



# The contribution of the body and motion to whole person recognition



Noa Simhi\*, Galit Yovel\*

*The Sagol School of Neuroscience, Tel-Aviv University, Tel-Aviv 69978, Israel*  
*The School of Psychological Sciences, Tel-Aviv University, Tel-Aviv 69978, Israel*

## ARTICLE INFO

### Article history:

Received 15 September 2015  
 Received in revised form 5 February 2016  
 Accepted 17 February 2016

### Keywords:

Person recognition  
 Face recognition  
 Body recognition  
 Biological motion  
 Form-from-motion processes  
 Dynamic identity signatures

## ABSTRACT

While the importance of faces in person recognition has been the subject of many studies, there are relatively few studies examining recognition of the whole person in motion even though this most closely resembles daily experience. Most studies examining the whole body in motion use point light displays, which have many advantages but are impoverished and unnatural compared to real life. To determine which factors are used when recognizing the whole person in motion we conducted two experiments using naturalistic videos. In Experiment 1 we used a matching task in which the first stimulus in each pair could either be a video or multiple still images from a video of the full body. The second stimulus, on which person recognition was performed, could be an image of either the full body or face alone. We found that the body contributed to person recognition beyond the face, but only after exposure to motion. Since person recognition was performed on still images, the contribution of motion to person recognition was mediated by form-from-motion processes. To assess whether dynamic identity signatures may also contribute to person recognition, in Experiment 2 we presented people in motion and examined person recognition from videos compared to still images. Results show that dynamic identity signatures did not contribute to person recognition beyond form-from-motion processes. We conclude that the face, body and form-from-motion processes all appear to play a role in unfamiliar person recognition, suggesting the importance of considering the whole body and motion when examining person perception.

© 2016 Elsevier Ltd. All rights reserved.

## 1. Introduction

When recognizing people in daily life we have a vast array of information at our disposal as we typically see the whole person and often in motion. However, research on the subject usually focuses on the role of faces in the process of person recognition. In the current study we focus on person recognition beyond the face and assess the roles of the body and motion in recognition of the whole person as well as the specific mechanisms involved in this process.

The processing of the whole body in motion has been primarily studied with point light displays (first described in Johansson, 1973). These displays are created by attaching reflective tape or LED lights to joints of the body and then filming a person while performing different types of actions. The resulting videos are then edited to create displays in which only the lights on the joints are visible and no other visual information. This enables the study of

\* Address: School of Psychological Sciences, Tel Aviv University, Tel Aviv 69978, Israel.

E-mail addresses: [simhinoa@gmail.com](mailto:simhinoa@gmail.com) (N. Simhi), [gality@post.tau.ac.il](mailto:gality@post.tau.ac.il) (G. Yovel).

motion with minimal contribution from other sources of information and such displays have been used to show that from point light motion it is possible to perceive the gender of a walker (Kozlowski & Cutting, 1977), their emotions (Atkinson, Dittrich, Gemmell, & Young, 2004), the weight of objects they are carrying (Runeson & Frykholm, 1983) and much more (for reviews see Blake & Shiffrar, 2007 and Thornton, 2006). Most relevant to the current study are studies showing person recognition based on point light displays of familiar or familiarized people (e.g. Cutting & Kozlowski, 1977; Hill & Pollick, 2000; Jacobs, Pinto, & Shiffrar, 2004; Loula, Prasad, Harber, & Shiffrar, 2005; Troje, Westhoff, & Lavrov, 2005), as well as matching of unfamiliar people (e.g. Richardson & Johnston, 2005; Stevenage, Nixon, & Vince, 1999). These studies varied greatly in the type of design and stimuli they employed as well as in the recognition rates they obtained, which varied from poor but above chance person recognition to sometimes near perfect recognition rates. Overall these findings suggest that motion can play a role in person recognition.

In naturalistic person recognition, isolating the role of motion is more complex, because while point light displays allow for the examination of motion independent of form information (if the

distance between the points of light is controlled for), in naturalistic situations we are also exposed to rich information from the face and body. Thus, studies using full light videos of people in motion are needed to isolate the roles of the face, body and motion and assess their relative contribution to person recognition. Several studies have used full light videos displaying people in motion for examining person recognition (e.g. Bruce, Henderson, Newman, & Burton, 2001; Burton, Wilson, Cowan, & Bruce, 1999; Liu, Seetzen, Burton, & Chaudhuri, 2003; Pilz, Vuong, Bülthoff, & Thornton, 2011; Roark, O'Toole, & Abdi, 2003; Roark, O'Toole, Abdi, & Barrett, 2006; Schiff, Banka, & de Bordes Galdi, 1986). In most of these studies, after seeing the full body in motion person recognition was examined from the face alone (Bruce et al., 2001; Burton et al., 1999; Liu et al., 2003; Pilz et al., 2011; Roark et al., 2006; Schiff et al., 1986) or, when person recognition was performed on the full body, only the face was presented at study (Roark, O'Toole, & Abdi, 2003). While many insights into face based person recognition have been gained using such paradigms, the contribution of the body to person recognition, beyond the face, cannot be assessed in these cases.

Recently several studies examined person recognition of the whole person in motion and highlighted very interesting interactions between the body and motion: The most comprehensive study on the subject was conducted by O'Toole et al. (2011). In their study, full body, face and body only stimuli were presented in pairs and participants had to determine if the same identity was presented in both stimuli. Each pair included the same type of images: either pairs of faces, pairs of full bodies or pairs of bodies alone. In order to examine the contribution of motion to person recognition, the two stimuli in each pair were presented in videos, multiple still images from the videos or a single static image from the video. Using this method O'Toole et al., 2011 showed that exposure to dynamic information improved person recognition in cases where the body was present, while motion did not contribute to person recognition based on the face alone. The same group recently examined the time course of person recognition in videos of familiarized people and highlighted the significance of the face and body to person recognition as it unfolds over time and viewing distance (Hahn, O'Toole, & Phillips, 2015). This study shows that the body contributes to person recognition at larger viewing distances from the observer, whereas at shorter distances the face dominates. Another recent study examined person recognition by presenting videos of the whole person and examining person recognition from either videos or static images of the full body, face or body alone (Robbins & Coltheart, 2015). This study found some advantages to full body compared to face recognition but no advantage to recognizing people from videos compared to still images. Overall, these studies show that the body and motion do contribute to person recognition beyond the face. Nevertheless, it cannot be determined what type of motion information plays a role in whole person recognition.

When examining the contribution of motion to whole person recognition, two processes should be taken into account: form-from-motion (or structure-from-motion) processes and dynamic identity signatures (described in O'Toole, Roark, & Abdi, 2002 in the context of faces). Form-from-motion processes refer to the creation of a better representation of the shape of a moving object relative to the representation of the same object's shape if viewed in static displays alone (e.g. see Koenderink, 1986 and Ullman, 1979 on the computational aspects that may be involved in these processes and how they might take place in human perception). Form-from-motion processes may play a role in person recognition for example if seeing a person in motion allows us to form a better representation of what that person's general body shape looks like when they stand still. Dynamic identity signatures on the other

hand refer to unique motion patterns that can be used for person recognition, or 'identity-specific' movements (O'Toole et al., 2002): a particular way of swinging one's hand or a slight limp for example can act as dynamic identity signatures and may be used as cues to recognize a person in motion (see Larsen, Simonsen, & Lynnerup, 2008 on the possible use of gait analysis in court).

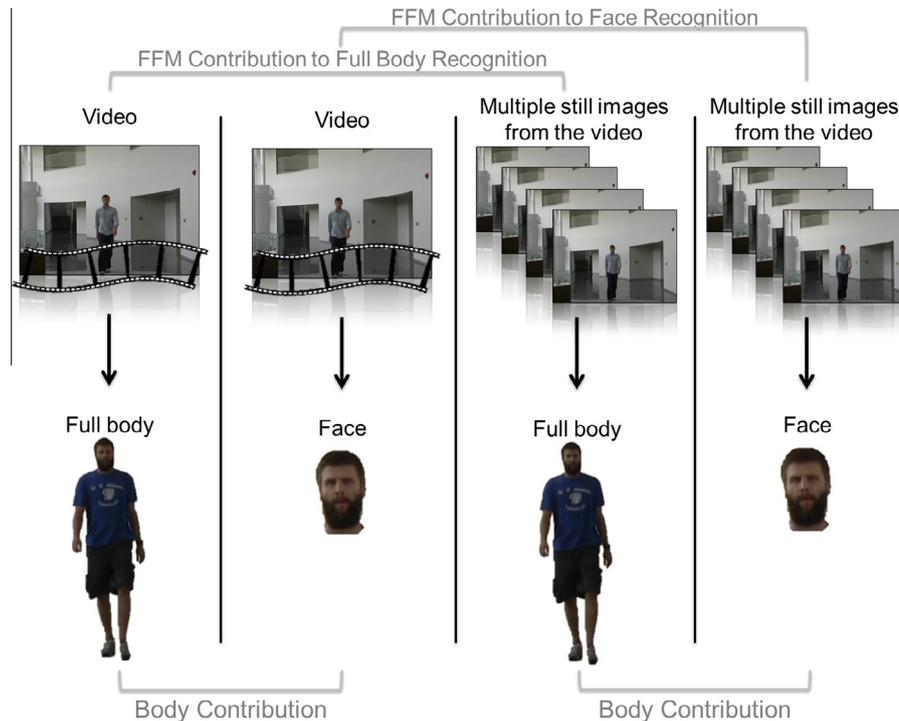
After seeing a person in motion both form-from-motion processes and dynamic identity signatures may play a role in person recognition. These processes can be dissociated by examining recognition from videos and still images separately. When recognizing a person from a still image for example, we rely mainly on form information since dynamic identity signatures are not available. When recognizing a person in video, dynamic identity signatures can contribute to person recognition beyond form-from-motion processes. Comparing person recognition between video and still images can therefore reveal the independent contribution of these two motion processes. It should be noted that motion may contribute to person recognition in additional ways (see for example theories on dynamic mental representations, Freyd, 1987), and the dissociation between these two processes may not be complete (as implied motion cues may be available in still images and dynamic information may better highlight form). However, the comparison between video and still images can provide some initial understanding of the relative roles of form and idiosyncratic motion related information in recognition of the whole dynamic person.

In the current study we thus outline whether the body and motion contribute to person recognition beyond the face, and examine the different types of motion processes that may mediate this contribution. In Experiment 1 we assessed the contribution of the body to person recognition and the role of form-from-motion processes. In Experiment 2 we examined whether dynamic identity signatures can contribute to whole person recognition beyond form-from-motion processes.

## 2. Experiment 1 – Do the face, body and form-from-motion processes contribute to person recognition?

To examine the relative contribution of the face, body and form-from-motion processes to person recognition we created a matching task in which the stimuli were presented sequentially, in pairs. The first stimulus in each pair always depicted the whole person and could either be a video clip of a person walking adapted from the Video Database of Moving Faces and People (O'Toole et al., 2005) or multiple still images from the video which contained similar visual information to that in the video but without motion. The second stimulus in the pair was a still image of a person depicting either the full body or face alone and participants were asked to determine if both of the stimuli in each pair depicted the same identity. Fig. 1 shows a schematic presentation of the experimental design.

Using this method the contribution of the body to person recognition could be assessed by examining the difference between full body and face based recognition. In particular, if person recognition from the full body proves better than person recognition from images of the face alone then the body contributes to person recognition beyond the face. The contribution of form-from-motion processes to person recognition could be assessed by examining the differences between recognition following exposure to videos vs. multiple still images. If person recognition from still images after exposure to videos proves better than person recognition after exposure to still images alone then form-from-motion processes contribute to person recognition.



**Fig. 1.** Schematic presentation of the general experimental design – in each trial participants were presented with either a video or multiple still images from the video and were asked to recognize the identity that was presented from a still image of either the full body or face which was taken on a different day (in the case of this example, the study and test stimuli depict the same person). The contribution of form-from-motion processes (FFM) is assessed by comparing performance between videos and still images. The contribution of the body is assessed by comparing recognition between the full body and face alone.

## 2.1. Methods

### 2.1.1. Participants

A total of 41 participants took part in the experiment (mean age = 24.41 years,  $SD = 2.05$ , 32 women) after being recruited at Tel Aviv University, either for course credit or payment. All participants had normal or corrected to normal vision and gave their informed consent to participate in the study by signing the appropriate consent form approved by the Tel Aviv University ethics committee and the experiment was performed in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki). Ten participants were included in each of the four experimental conditions described below. One participant was removed from analysis due to significantly below chance performance (22% correct responses) and was replaced.

### 2.1.2. Stimuli

Forty-five identities were selected from the Database of Moving Faces and People (O'Toole et al., 2005) (15 men). For each identity two videos filmed on different occasions (which could be up to six months apart) were selected for stimuli creation, depicting participants walking towards the camera at a normal pace in a naturally lit corridor. The earliest video of the two was used to create the study stimulus in each pair (which could either be a video or multiple still images) and the second video was used to create the test stimulus (a still image of the full body or face alone).

The study stimuli were created by first cutting all of the videos so that the video ended when the participants' shoes were no longer visible so that the full body was displayed in every frame of the video. These videos were 2–7 s in length (with a mean length of 5.6 s) and were roughly  $18.3^\circ$  by  $12.3^\circ$  in visual angle during the experiment ( $720 \times 480$  pixels on the screen). The varying length of the videos reflects natural differences in the participants' walking

pace and was maintained in all the different experimental conditions.

To create the multiple still image displays, one frame per each second of the video was extracted (see Fig. 2 for an example), with the last image always matching the last frame of the video. During the experiment these images were presented in the order in which they appeared in the video with a 0.2 s inter stimulus interval in order to match the duration of the videos as closely as possible while minimizing the impression of apparent motion. It should be noted however that even if apparent motion perception occurred during these study conditions, any differences between the study of videos and the study of multiple still images would only further highlight the importance of viewing actual motion in person recognition. The images, like the videos, were roughly  $18.3^\circ$  by  $12.3^\circ$  in visual angle ( $720 \times 480$  pixels on the screen) and the number of images presented varied in accordance with the length of the video they were extracted from.

The test stimulus in each pair was extracted from the second video of each identity. For each identity two types of test stimuli were created (depicted in Fig. 1):

- Full body images** – these images were created from the latest frame in the original video where the full person was visible and in as neutral a pose as possible. These frames were edited using the Microsoft Office (Microsoft Corp.) background removal tool so that no background information remained and were then resized to take up roughly the same area on the screen,  $3.3^\circ$  by  $8.4^\circ$  visual angle ( $130 \times 330$  pixels), these dimensions were allowed to vary somewhat however, reflecting individual differences such as weight and height.
- Face images** – these images were extracted from the same original videos as the full body images, but from the frame containing the best view of the face. This frame generally appeared after the frame used to create the full body images



Fig. 2. Multiple frames extracted from a video and used for the multiple still stimuli, as detailed in the stimuli Section 2.1.2.

since in the original, unedited, videos participants moved close to the camera towards the end of the video so that the full body was no longer visible but the face was viewed from a relatively short distance. The background was then removed from these frames and the images cropped so that only the face and neck were visible. Finally the stimuli were resized to take up roughly the same area on screen,  $3.3^\circ$  by  $3.8^\circ$  visual angle ( $130 \times 150$  pixels) allowing for some variance, as above, reflecting individual differences in face shape and size.

### 2.1.3. Stimuli pairing

Due to the large variability in physical appearance of the identities in the Video Database of Moving Faces and People (O'Toole et al., 2005), even when only Caucasians are selected (as was the case in this study in order to avoid other race effects), identities were divided into groups according to similarity in visual appearance. These groups were used to create the mismatch pairs in the experiment so that the pairings were not exceedingly easy (this ensured that male and female identities as well as identities with visible differences in height, skin tone, weight etc. were not matched together).

Separate groups were created based on similarity in full body images and similarity in facial images since people who were similar in full body images were not necessarily similar in images of the face alone. Similarity was determined independently by two experimenters who then settled disputes together, with the assistance of a third observer in the few cases when no agreement could be reached. This resulted in fifteen similarity groups in each of the test conditions containing 44 identities in total (as one of the 45 identities in the experiment was used in a practice trial): so there were fourteen groups containing three similarly looking identities and one group with two similarly looking identities.

In order to make sure that the stimuli pairings in the full body and face conditions did not differ in their performance level, in a preliminary study we tested two groups of subjects ( $n = 10$  in each group) where one group performed the experiment with the pairings based on full body similarity and the other with the pairings based on face similarity. We used  $d'$  to assess performance level

(see Section 2.1.6). An independent sample  $t$ -test revealed no difference in performance between the two groups ( $t(18) = .79$ ,  $p = .44$ , Cohen's  $d = .35$ ). This indicates that any difference between the full body and face test conditions in the experiment cannot be attributed to a difference in the type of mismatched pairs that were used in the experiment.

### 2.1.4. Design

The experiment constituted of a matching task in which the mode of presentation of the study stimuli (videos or multiple still images) and the type of test stimuli (full body or face images) both varied between participants so that each participant saw only one type of study stimuli and one type of test stimuli (as depicted in Fig. 1).

Half of the trials in each of the conditions were match trials and half were mismatch trials, with the mismatch trial pairings selected randomly from within groups of similar identities (as described in the stimuli Section 2.1.2.).

### 2.1.5. Apparatus and procedure

The experiment was presented in MATLAB using the Psychophysics Toolbox extensions (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997) on a Samsung SyncMaster SA950, Full HD, LED monitor with a  $1920 \times 1080$  screen resolution, in front of which the participants were seated at a comfortable distance of approximately 60 cm.

During the experiment participants were presented with two stimuli in succession and asked to determine if both of the stimuli depicted the same or different identities, as described below.

Before the experiment commenced participants were presented with one practice trial in order to ensure that they understood the task. During the practice trial participants were informed that the first and second stimuli in each pair that they would be presented with were not filmed on the same day and therefore changes in the appearance of the people filmed (for example their clothing and hair style) were to be expected. Performance on this trial was not recorded or used for analysis, and stimuli in this trial were of the same type as those in the following experimental trials.

In the practice trial and each of the 44 experimental trials that followed, participants were presented with a fixation point which appeared for 0.75 s at the center of the screen and was replaced by a study stimulus. 0.5 s after the study stimulus a fixation point appeared once again for 0.75 s and was replaced by a test stimulus. During the presentation of the test stimulus participants were asked to respond using 'yes' and 'no' keys whether the test stimulus depicted the same identity as the study stimulus. The test stimulus remained on screen until the participants responded, at which point the screen was replaced by a confidence scale ranging from 1 (lowest confidence) to 5 (highest confidence) and participants were asked to indicate using the appropriate keys how confident they were in the decision they just made. After this response a fixation point appeared once again for 0.75 s and the next study stimulus was presented. See Fig. 3 for a summary of the time course of each trial.

Every ten trials participants took a mandatory break of 5 s at least. Participants were asked to keep their fingers on the 'yes' and 'no' response keys for the duration of the experiment (which generally took less than 15 min).

### 2.1.6. Data analysis

Participants performance level was assessed using the measure of sensitivity,  $d'$ . In calculating the  $d'$  hits were defined as match trials in which the participants responded 'yes', indicating that they believed that the two stimuli depicted the same identity, and false alarms were defined as mismatch trials in which the participants responded 'yes' despite the identities being different.

We also looked at the participants' response bias, as measured by the criterion,  $C$ , where a positive value indicates a conservative response bias and a tendency to respond that the two stimuli depict different identities and a negative value indicates a liberal response bias and a tendency to respond that the stimuli depict

the same identity. Finally, we also analyzed participants' confidence ratings on correct trials in the different experimental conditions.

Statistical analysis was performed using Statistica 9.0 (StatSoft Inc).

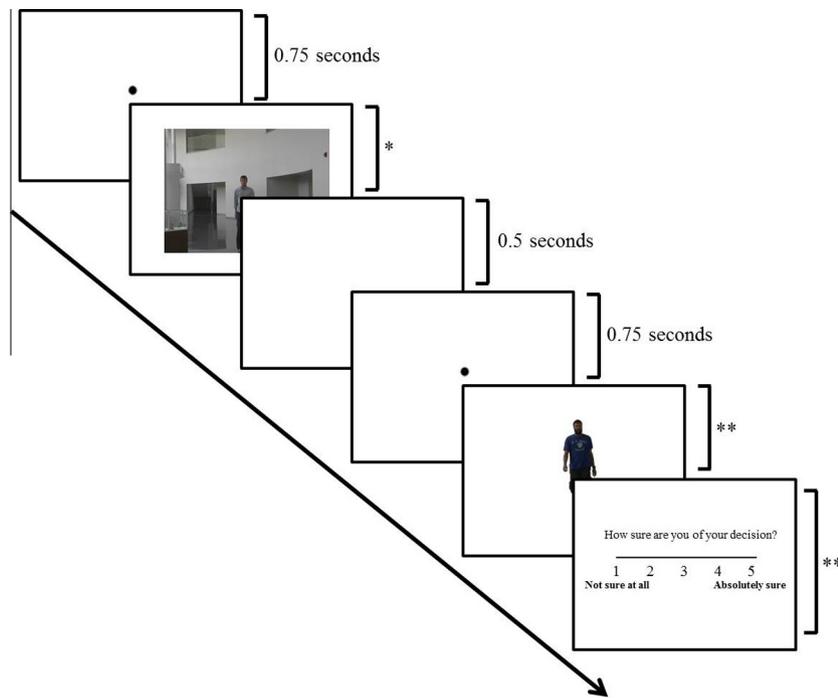
### 2.2. Results

A factorial  $2 \times 2$  ANOVA with Study Stimulus (Video, Multiple stills images) and Test Stimulus (Full body, Face) as between subject factors was performed on the  $d'$  of participants in each group and revealed a main effect of Study Stimulus ( $F(1,36) = 5.65$ ,  $p = .023$ ,  $\eta_p^2 = .14$ ), a main effect of Test Stimulus ( $F(1,36) = 5.24$ ,  $p = .028$ ,  $\eta_p^2 = .13$ ) and an interaction ( $F(1,36) = 5.54$ ,  $p = .024$ ,  $\eta_p^2 = .13$ ).

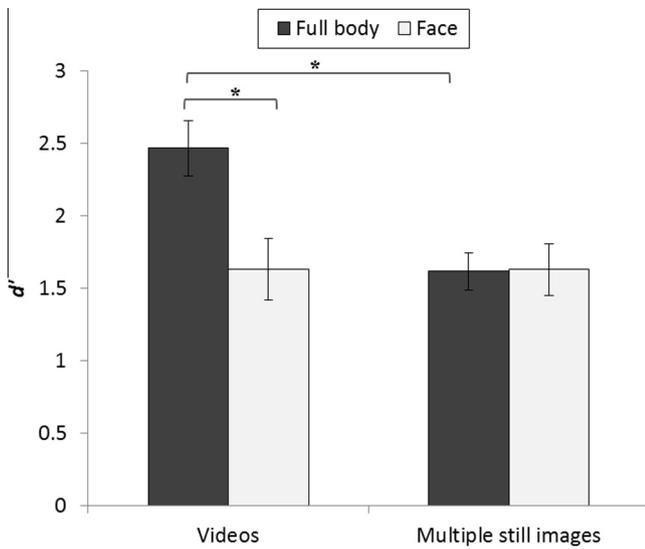
Post-hoc Tukey HSD analysis of the interaction revealed that person recognition from the full body was superior to person recognition from the face alone after studying videos ( $p = 0.012$ ) but person recognition from the full body and face did not differ when studying still images alone ( $p = 0.99$ ). Furthermore, while person recognition from the full body was better after the study of videos compared to multiple still images ( $p = 0.01$ ) person recognition from the face alone did not differ between study conditions ( $p = 0.99$ ). Fig. 4 depicts these results.

Analysis of response bias as measured by criterion,  $C$ , revealed no main effect of Study Stimulus ( $F(1,36) = .12$ ,  $p = .913$ ,  $\eta_p^2 = .0003$ ), a trend towards an effect of Test Stimulus ( $F(1,36) = 3.23$ ,  $p = .08$ ,  $\eta_p^2 = .08$ ) due to a higher criterion in person recognition from the face compared to person recognition from the full body, and no interaction ( $F(1,36) = .51$ ,  $p = .478$ ,  $\eta_p^2 = .01$ ).

Analysis of the confidence ratings on correct trials revealed no main effect of Study Stimulus ( $F(1,36) = .23$ ,  $p = .64$ ,  $\eta_p^2 = .006$ ), Test Stimulus ( $F(1,36) = 1.21$ ,  $p = .28$ ,  $\eta_p^2 = .03$ ) or interaction ( $F(1,36)$



**Fig. 3.** The time course of a given trial: Each trial started with a fixation point presented for 0.75 s followed by a study stimulus (which could be either multiple still images or a video and varied between participants). The study stimulus was replaced by a blank screen which was presented for 0.5 s, after which time a fixation point was once again presented for 0.75 s followed by a test stimulus (which varied between participants and could be either a full body or face image) which was presented until response. After response the screen was replaced by a confidence scale and participants were asked to indicate on a scale of 1–5 how confident they were in their decision (the text during this part of the experiment appeared in Hebrew and is translated here for convenience). \*Study stimulus duration varied between trials, as detailed in the stimuli Section 2.1.2. \*\*Presented until response.



**Fig. 4.** The  $d'$  of person recognition when the study stimuli were videos or multiple still images, and the test stimuli were full body or face images. Error bars indicate the standard error of the mean. \*  $p < .05$ .

= 1.64,  $p = .21$ ,  $\eta_p^2 = .04$ ). The mean confidence ratings, criterion (C) values and hit and false alarm rates in the experiment can be seen in Table 1.

2.3. Discussion

Experiment 1 reveals two important findings. First, the body contributes to person recognition beyond the face only when we are exposed to people in motion. Second whole person recognition is better following exposure to moving than static images of full bodies. We first discuss the contribution of the body beyond the face to person recognition and then the role of form-from-motion processes in whole person recognition.

Our findings show superior person recognition from the full body compared to the face alone, but only after exposure to the person in motion. The finding that after exposure to still images alone person recognition from the full body was no different than person recognition from the face alone is particularly striking, indicating that without previous information from motion the body does not contribute to person recognition beyond the face. These findings are consistent with O'Toole et al. (2011), who also reported in a within media matching task no effect of motion on face recognition and an advantage to dynamic displays when the body is present. Here we extend these findings in two ways: first, we show that after exposure to motion the body can contribute to

**Table 1**

The mean and standard error of the mean (SEM) of the hit rates, false alarm rates, criterion (C) and confidence ratings on correct trials in each of the experimental conditions in Experiment 1.

Study stimulus	Videos		Multiple still images	
	Full body	Face	Full body	Face
Hits	.87	.75	.80	.72
SEM	.03	.03	.03	.03
False Alarms	.13	.22	.25	.18
SEM	.03	.05	.04	.04
Criterion (C)	-.02	.10	-.09	.19
SEM	.11	.11	.12	.09
Confidence ratings on correct trials	4.11	3.8	4.01	4.03
SEM	.11	.17	.11	.13

person recognition from still images as well, indicating the role of form-from-motion processes in whole person recognition. Second, we apply these findings to recognition of the whole person and show that recognition of the whole person previously seen in motion is based on both the body and the face whereas bodies do not contribute to person recognition beyond the face when the whole person was previously seen in static images.

It is important to note that viewing the full body at test did not produce a general advantage to person recognition, and so the possibility that the full body test images may have been more similar to the full body study videos does not explain our results. Similarly, the additional information that is inherently contained in moving stimuli compared to a few still images taken from the video did not produce a general advantage to person recognition, as person recognition based on the face alone did not differ following video or multiple still image presentations. Our results are also incompatible with the possibility of a greater cost to generalizing between media (i.e., a video and still image) than within media (multiple still images to a single still image), since performance was never better when studying multiple still images compared to videos.

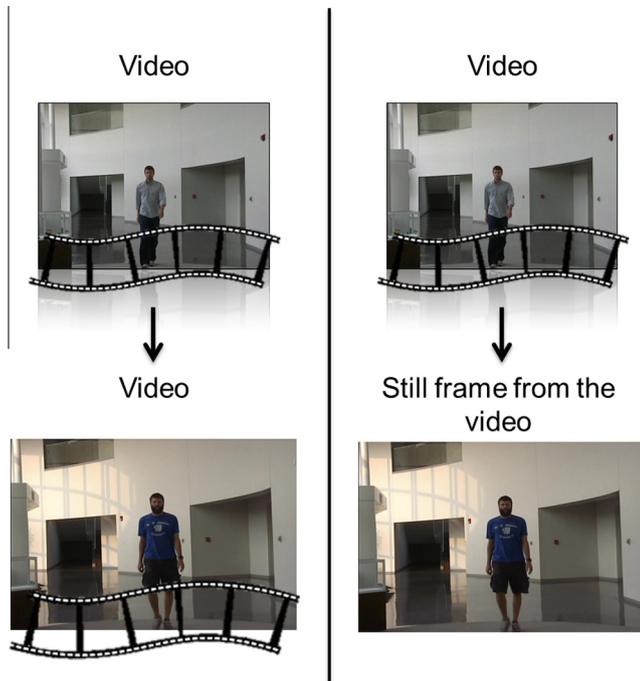
It should also be noted that by employing a between-subjects design, person recognition from the face alone was examined in ideal conditions. Since participants in the face group knew that face only images would always appear at test, they could have focused on the face alone when viewing the full body videos. Thus, the advantage to full body compared to face only recognition may be even greater in a within-subject design when full body and face images at test are randomly presented in the same design and therefore participants who do not know in advance if the body will be present at test or not, need to study the whole person on all trials. This prediction may be studied in future experiments.

A second important finding of Experiment 1 was that full body recognition was better following the presentation of moving people than multiple still images of these people. Given that person recognition in our study was always performed from still images alone, and not from videos, the body's contribution to person recognition after exposure to motion cannot be mediated by dynamic identity signatures and was therefore enabled by form-from-motion processes. Thus, exposure to the person in motion improved the representation of that person's general physical appearance and thereby improved person recognition even from images that were taken on a different day.

To further explore the mechanisms underlying the role of motion in person recognition, in Experiment 2 we examined whether dynamic identity signatures contribute to person recognition beyond form-from-motion processes. To that end, we compared person recognition from the full body in videos and still images after exposure to people in motion (see Fig. 5). Using this comparison we could determine if the availability of dynamic identity signatures in person recognition from videos improves person recognition beyond form-from-motion processes, that is person recognition from still images alone. The combination of the two experiments in this study thus allows us to separately assess the contributions of form-from-motion processes and dynamic identity signatures in person recognition.

**3. Experiment 2 – Do dynamic identity signatures contribute to person recognition?**

To examine the contribution of dynamic identity signatures in person recognition, in this experiment the study stimuli were videos, as before, however test stimuli could either be a single frame from the video or the video itself. An advantage to video



**Fig. 5.** A schematic presentation of the design in Experiment 2. In each trial participants were presented with a video and were asked to recognize the identity that was presented in that video from either a video which was filmed on a different day or a still frame taken from the video.

compared to image based person recognition in this case would suggest a contribution of dynamic identity signatures.

### 3.1. Method

#### 3.1.1. Participants

20 participants took part in this experiment (mean age = 23.85 years,  $SD = 2.25$ , 14 women) after being recruited at Tel Aviv University, either for course credit or payment. All participants had normal or corrected to normal vision and gave their informed consent to participate in the study by signing the appropriate consent form approved by the Tel Aviv University ethics committee and the experiment was performed in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki). Ten participants were randomly allocated to each of the experimental conditions.

#### 3.1.2. Stimuli

The study videos from Experiment 1 were used as the study stimuli in this experiment as well. The test stimuli were also created from the same videos from which the test stimuli in Experiment 1 were extracted but in this case could be either:

- Videos** – in order to create these stimuli the original videos were cut so that each of them was 4 s in length and always ended at the point in which the participants' shoes were no longer visible. These videos took up roughly  $18.3^\circ$  by  $12.3^\circ$  visual angle ( $720 \times 480$  pixels on the screen).
- Still frames** – these images consisted of the final frame in the test videos described above unless the video ended mid-stride in which case the most recent frame in which the participant was in as neutral a pose as possible was selected. Note that these differed from the full body stimuli in Experiment 1 mainly in that the background was now included in the test image in order to match the video test condition as closely as possible.

#### 3.1.3. Design

The design in Experiment 2 was the same as in Experiment 1. The similarity groups that were used in order to create the mismatch pairs were based on the full body images in both test conditions (see Section 2.1.3 for details on stimulus pairing). See Fig. 5 for a schematic presentation of the design.

#### 3.1.4. Apparatus and procedure

The apparatus and procedure were the same as described in Experiment 1 except that the test stimulus in each pair of stimuli in this case (be it a video or a single still frame) appeared for 4 s only and then was replaced by a blank screen which was displayed until the participant responded unless a response was received during the stimulus presentation itself, in which case the next trial sequence began.

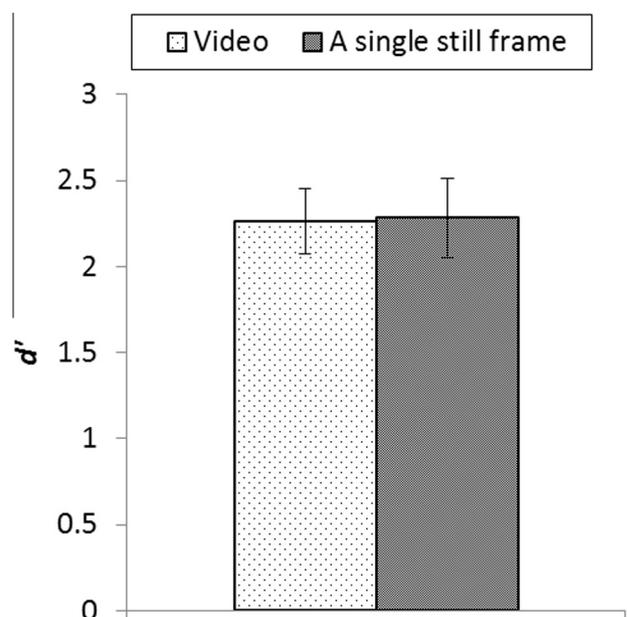
### 3.2. Results

An independent samples *t*-test revealed no differences in performance as measured by  $d'$  between person recognition from video and person recognition from a still frame from the video following exposure to the person in motion ( $t(18) = .07$ ,  $p = .94$ , Cohen's  $d = .03$ ). This can be seen in Fig. 6.

In addition an independent samples *t*-test did not reveal differences in response bias as measured by the criterion,  $C$ , between the two test conditions in Experiment 2 ( $t(18) = .02$ ,  $p = .98$ , Cohen's  $d = .01$ ) and no differences were found in the confidence ratings on correct trials in the two conditions ( $t(18) = .6$ ,  $p = .56$ , Cohen's  $d = .27$ ). Mean values for these measures as well as the hit rates and false alarms can be found in Table 2.

### 3.3. Discussion

In Experiment 2 we found that person recognition from a single still frame from a video was no different than recognition from a video itself after the person was seen in motion. This indicates that in the current experiment dynamic identity signatures did not contribute to person recognition beyond the form-from-motion



**Fig. 6.** The  $d'$  of person recognition in Experiment 2. In Experiment 2 the study stimulus was always a video and the test stimulus was either a video or a still frame from the video (see Fig. 5 for a schematic presentation of this design). Error bars indicate the standard error of the mean.

**Table 2**

The mean and standard error of the mean of the hit rates, false alarm rates, criterion (C) and confidence ratings on correct trials in each of the experimental conditions in Experiment 2.

Study stimulus	Video	
	Video	Still frame from the video
Hits	.83	.83
SEM	.03	.03
False Alarms	.14	.16
SEM	.04	.05
Criterion (C)	.10	.09
SEM	.13	.16
Confidence on correct trials	4.03	4.15
SEM	.12	.15

processes that we revealed in Experiment 1. Whereas performance on this task was relatively high (mean 84.2% correct responses,  $SD = 7.8$ ) there was still considerable room for improvement if dynamic identity signatures can be used for person recognition and so it is unlikely that these results are due to a ceiling effect. Furthermore, our findings are consistent with Robbins and Coltheart (2015), who also revealed no advantage to recognition from videos compared to still images after viewing the whole person in motion, again suggesting that dynamic identity signatures were not used for recognition. It should be noted that in both ours and the Robbins & Coltheart, 2015 study, performance was similar across recognition conditions despite the fact that one of them presented same media stimuli in study and test (video–video) and the second different media stimuli (video–still). Thus, generalization across vs. within media does not account for these findings.

The lack of an advantage to person recognition from video may be surprising in light of the research available from point light displays in which above chance person recognition has been demonstrated (e.g. Cutting & Kozlowski, 1977; Loula et al., 2005). In such studies mainly dynamic identity signatures likely contribute to person recognition even though some remnants of form cues may still be available. Several factors may account for this apparent discrepancy. First, dynamic identity signatures may be used in person recognition only when little form cues are available, such as in point light displays, while in conditions where the information from form is less ambiguous it is the primary contributor to person recognition with little additional contribution to motion cues in themselves. Second, some dynamic identity signatures may be more informative than others. For example, in Loula et al. (2005) it was shown that person recognition from point light displays of dancing, boxing, ping-pong playing or jumping were better for person recognition than displays of people walking. It is possible therefore that the type of action performed influences how informative dynamic identity signatures prove to be in person recognition. Third, person recognition in our experiment was performed on unfamiliar people after a single exposure. It is possible that dynamic identity signatures require more experience in order to contribute to person recognition (indeed many point light studies are conducted on familiar or familiarized identities) and that while dynamic identity signatures do not appear to play a role in unfamiliar person recognition they may improve familiar person recognition (as suggested in the context of facial recognition in O'Toole et al. (2002)).

Finally, it should be noted that one aspect of motion information that we did not consider in the current study is implied motion that may be extracted from static images of the whole person. Indeed, images that convey implied motion have been shown to activate motion-selective neural mechanisms (e.g., Kourtzi & Kanwisher, 2000), but their role in person recognition is still unknown. To assess the role of implied motion, person recognition from static images that convey dynamic or static information should be compared in future studies.

## 4. General discussion

In this study we conducted two experiments examining the role of the face, body and motion in person recognition. We conclude that the body contributes to person recognition beyond the face but only after exposure to the person in motion. Furthermore, the body's contribution to person recognition was enabled by form-from-motion processes, while dynamic identity signatures appear to play no role in our task.

To the best of our knowledge, this is the first study in which the independent contributions of form-from-motion processes and dynamic identity signatures have been assessed in the context of whole person recognition and within the same study. The distinction between from-from-motion processes (i.e. the representational enhancement hypothesis) and dynamic identity signatures (i.e., the supplemental information hypothesis) has been originally discussed in the context of face recognition (as reviewed in O'Toole et al., 2002; O'Toole & Roark, 2010; Xiao et al., 2014). Overall the effect of motion on face recognition varies across different conditions. For example, non-rigid motion (e.g., expression, speaking) has been shown to enhance face recognition more than rigid motion (e.g. head rotation) (Lander & Bruce, 2003, but see Christie & Bruce, 1998). Dynamic identity signatures are usually considered to play a larger role in familiar person recognition or in cases of uncertainty (O'Toole & Roark, 2010; O'Toole et al., 2002; Xiao et al., 2014). Here we expand the discussion on the role of motion beyond faces and show its importance in recognition of the whole person.

Our results, adding onto the recent studies by Hahn, O'Toole, and Phillips (2015), O'Toole et al. (2011), Robbins and Coltheart (2015), further stress the importance of examining the person as a whole, including both the face and body, in the context of person recognition. Indeed, even though relatively few studies have examined the whole person and not the face alone, these studies show interesting and promising results with respect to the role the body plays beyond the face. In eye tracking studies the body has been shown to mediate gazes to the face in person detection (Bindemann, Scheepers, Ferguson, & Burton, 2010). Bodies have also been shown to play an important role in emotion recognition (Aviezer, Trope, & Todorov, 2012) and even face based person recognition has been shown to be improved when viewing faces in the context of an approaching 3D body model (Pütz et al., 2011).

As mentioned, the importance of the body and face in person recognition as it unfolds over time was further examined in a recent comprehensive study (Hahn et al., 2015) in which recognition of familiarized identities was assessed from videos of people approaching. Here the video length that was presented during person recognition was varied as well as the information available for recognition, which could either be the face, body or full body. The authors find that the body made an independent contribution to person recognition only when people were recognized from a distance, whereas from closer distances recognition relied predominantly on the face. In a study by Burton et al. (1999) on familiar person recognition, it was shown that although the face was the primary contributor to person recognition from video, the gait and body contributed as well.

Studies that examined the role of the body with static images have also found above chance person recognition based on the body alone even though the face was the primary contributor to person recognition (Robbins & Coltheart, 2012). Furthermore, in cases where information from the face is misleading or unreliable the body can be used in person recognition even if the participants are unaware of doing so (Rice, Phillips, Natu, An, & O'Toole, 2013). It has further been suggested that in poor viewing conditions person recognition from the body and face may be equally accurate (Rice, Phillips, & O'Toole, 2013).

Taken together these studies suggest that the body and motion contribute to person recognition in a variety of ways and that in conditions of uncertainty they may be especially important. Interesting directions for future studies include examining how the contribution of dynamic identity signatures to person recognition might be affected by familiarity as well as examining how the contribution of the face, body and motion might be influenced by different retention intervals in memory, which is more similar to person recognition in real life when we need to recognize previously seen people across longer periods of time. Finally, these findings may also have implications in regards to the design of suspect recognition line ups where the common practice is to make use of photo line ups including pictures of the face alone while the contribution of the body and form-from-motion processes found in this study suggests that the use of full body images may improve suspect recognition in such cases, a possibility which merits further investigation.

## Acknowledgments

We would like to thank Alice O'Toole for providing us with the video database of moving faces and people. This work was supported by an Israeli Science Foundation (ISF) grant 446/12 to GY. Special thanks to undergraduate students Ofir Becker and Rita Krigman who assisted in stimuli preparation, participant recruitment and data gathering.

## References

- Atkinson, A. P., Dittrich, W. H., Gemmell, A. J., & Young, A. W. (2004). Emotion perception from dynamic and static body expressions in point-light and full-light displays. *Perception*, 33(6), 717–746. <http://dx.doi.org/10.1068/p5096>.
- Aviezer, H., Trope, Y., & Todorov, A. (2012). Body cues, not facial expressions, discriminate between intense positive and negative emotions. *Science (New York, N.Y.)*, 338(6111), 1225–1229. <http://dx.doi.org/10.1126/science.1224313>.
- Bindemann, M., Scheepers, C., Ferguson, H. J., & Burton, A. M. (2010). Face, body, and center of gravity mediate person detection in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance*, 36(6), 1477–1485. <http://dx.doi.org/10.1037/a0019057>.
- Blake, R., & Shiffrar, M. (2007). Perception of human motion. *Annual Review of Psychology*, 58, 47–73. <http://dx.doi.org/10.1146/annurev.psych.57.102904.190152>.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10(4), 433–436. <http://dx.doi.org/10.1163/156856897X00357>.
- Bruce, V., Henderson, Z., Newman, C., & Burton, A. M. (2001). Matching identities of familiar and unfamiliar faces caught on CCTV images. *Journal of Experimental Psychology: Applied*, 7(3), 207–218. <http://dx.doi.org/10.1037/1076-898X.7.3.207>.
- Burton, A. M., Wilson, S., Cowan, M., & Bruce, V. (1999). Face recognition in poor-quality video: evidence from security surveillance. *Psychological Science*, 10(3), 243–248. <http://dx.doi.org/10.1111/1467-9280.00144>.
- Christie, F., & Bruce, V. (1998). The role of dynamic information in the recognition of unfamiliar faces. *Memory & Cognition*, 26(4), 780–790. <http://dx.doi.org/10.3758/BF03211397>.
- Cutting, J. E., & Kozlowski, L. T. (1977). Recognizing friends by their walk: Gait perception without familiarity cues. *Bulletin of the Psychonomic Society*, 9(5), 353–356. <http://dx.doi.org/10.3758/BF03337021>.
- Freyd, J. J. (1987). Dynamic mental representations. *Psychological Review*, 94(4), 427–438. <http://dx.doi.org/10.1037/0033-295X.94.4.427>.
- Hahn, C. A., O'Toole, A. J., & Phillips, P. J. (2015). Dissecting the time course of person recognition in natural viewing environments. *British Journal of Psychology*, 1–18. <http://dx.doi.org/10.1111/bjop.12125> (London, England: 1953).
- Hill, H., & Pollick, F. E. (2000). Exaggerating temporal differences enhances recognition of individuals from point light displays. *Psychological Science*, 11(3), 223–228. <http://dx.doi.org/10.1111/1467-9280.00245>.
- Jacobs, A., Pinto, J., & Shiffrar, M. (2004). Experience, context, and the visual perception of human movement. *Journal of Experimental Psychology: Human Perception and Performance*, 30(5), 822–835. <http://dx.doi.org/10.1037/0096-1523.30.5.822>.
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, 14(2), 201–211. <http://dx.doi.org/10.3758/BF03212378>.
- Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Brissard, C. (2007). What's new in psychtoolbox-3. *Perception*, 36(14), 1.
- Koenderink, J. J. (1986). Optic flow. *Vision Research*, 26(1), 161–179. [http://dx.doi.org/10.1016/0042-6989\(86\)90078-7](http://dx.doi.org/10.1016/0042-6989(86)90078-7).
- Kourtzi, Z., & Kanwisher, N. (2000). Activation in human MT/MST by static images with implied motion. *Journal of Cognitive Neuroscience*, 12(1), 48–55. <http://dx.doi.org/10.1162/08989290051137594>.
- Kozlowski, L. T., & Cutting, J. E. (1977). Recognizing the sex of a walker from a dynamic point-light display. *Perception & Psychophysics*, 21(6), 575–580. <http://dx.doi.org/10.3758/BF03198740>.
- Lander, K., & Bruce, V. (2003). The role of motion in learning new faces. *Visual Cognition*, 10(8), 897–912. <http://dx.doi.org/10.1080/13506280344000149>.
- Larsen, P. K., Simonsen, E. B., & Lynnerup, N. (2008). Gait analysis in forensic medicine. *Journal of Forensic Sciences*, 53(5), 1149–1153. <http://dx.doi.org/10.1111/j.1556-4029.2008.00807.x>.
- Liu, C. H., Seetzen, H., Burton, A. M., & Chaudhuri, A. (2003). Face recognition is robust with incongruent image resolution: Relationship to security video images. *Journal of Experimental Psychology: Applied*, 9(1), 33–41. <http://dx.doi.org/10.1037/1076-898X.9.1.33>.
- Loula, F., Prasad, S., Harber, K., & Shiffrar, M. (2005). Recognizing people from their movement. *Journal of Experimental Psychology: Human Perception and Performance*, 31(1), 210–220. <http://dx.doi.org/10.1037/0096-1523.31.1.210>.
- O'Toole, A. J., Harms, J., Snow, S. L., Hurst, D. R., Pappas, M. R., Ayyad, J. H., et al. (2005). A video database of moving faces and people. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(5), 812–816. <http://dx.doi.org/10.1109/TPAMI.2005.90>.
- O'Toole, A. J., Phillips, P. J., Weimer, S., Roark, D. A., Ayyad, J., Barwick, R., et al. (2011). Recognizing people from dynamic and static faces and bodies: dissecting identity with a fusion approach. *Vision Research*, 51(1), 74–83. <http://dx.doi.org/10.1016/j.visres.2010.09.035>.
- O'Toole, A. J., Roark, D. A., & Abdi, H. (2002). Recognizing moving faces: A psychological and neural synthesis. *Trends in Cognitive Sciences*, 6(6), 261–266. [http://dx.doi.org/10.1016/S1364-6613\(02\)01908-3](http://dx.doi.org/10.1016/S1364-6613(02)01908-3).
- O'Toole, A., & Roark, D. (2010). *Memory for moving faces: The interplay of two recognition systems*. In *Dynamic Faces: The MIT Press*. 10.7551/mitpress/9780262014533.003.0003.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, 10(4), 437–442. <http://dx.doi.org/10.1163/156856897X00366>.
- Pilz, K. S., Vuong, Q. C., Bülhoff, H. H., & Thornton, I. M. (2011). Walk this way: Approaching bodies can influence the processing of faces. *Cognition*, 118(1), 17–31. <http://dx.doi.org/10.1016/j.cognition.2010.09.004>.
- Rice, A., Phillips, P. J., Natu, V., An, X., & O'Toole, A. J. (2013). Unaware person recognition from the body when face identification fails. *Psychological Science*, 24(11), 2235–2243. <http://dx.doi.org/10.1177/0956797613492986>.
- Rice, A., Phillips, P. J., & O'Toole, A. (2013). The role of the face and body in unfamiliar person identification. *Applied Cognitive Psychology*, 27(6), 761–768. <http://dx.doi.org/10.1002/acp.2969>.
- Richardson, M. J., & Johnston, L. (2005). Person recognition from dynamic events: The kinematic specification of individual identity in walking style. *Journal of Nonverbal Behavior*, 29(1), 25–44. <http://dx.doi.org/10.1007/s10919-004-0888-9>.
- Roark, D. A., O'Toole, A. J., & Abdi, H. (2003). Human recognition of familiar and unfamiliar people in naturalistic video. In *2003 IEEE International SOI Conference. Proceedings (Cat. No.03CH37443)* (pp. 36–41). IEEE Comput. Soc. <http://dx.doi.org/10.1109/AMFG.2003.1240821>.
- Roark, D. A., O'Toole, A. J., Abdi, H., & Barrett, S. E. (2006). Learning the moves: The effect of familiarity and facial motion on person recognition across large changes in viewing format. *Perception*, 35(6), 761–773. <http://dx.doi.org/10.1068/p5503>.
- Robbins, R. A., & Coltheart, M. (2012). The effects of inversion and familiarity on face versus body cues to person recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 38(5), 1098–1104. <http://dx.doi.org/10.1037/a0028584>.
- Robbins, R. A., & Coltheart, M. (2015). The relative importance of heads, bodies, and movement to person recognition across development. *Journal of Experimental Child Psychology*, 138, 1–14. <http://dx.doi.org/10.1016/j.jecp.2015.04.006>.
- Runeson, S., & Frykholm, G. (1983). Kinematic specification of dynamics as an informational basis for person-and-action perception: Expectation, gender recognition, and deceptive intention. *Journal of Experimental Psychology: General*, 112(4), 585–615. <http://dx.doi.org/10.1037/0096-3445.112.4.585>.
- Schiff, W., Banka, L., & de Bordes Galdi, G. (1986). Recognizing people seen in events via dynamic "Mug Shots". *The American Journal of Psychology*, 99(2), 219. <http://dx.doi.org/10.2307/1422276>.
- Stevenage, S. V., Nixon, M. S., & Vince, K. (1999). Visual analysis of gait as a cue to identity. *Applied Cognitive Psychology*, 13(6), 513–526. [http://dx.doi.org/10.1002/\(SICI\)1099-0720\(199912\)13:6<513::AID-ACP16>3.0.CO;2-8](http://dx.doi.org/10.1002/(SICI)1099-0720(199912)13:6<513::AID-ACP16>3.0.CO;2-8).
- Thornton, I. M. (2006). Biological motion: Point-light walkers and beyond. In G. Knoblich, I. M. Thornton, M. Grosjean, & M. Shiffrar (Eds.), *Human body perception from the inside out* (pp. 271–305). New York: Oxford University Press.
- Troje, N. F., Westhoff, C., & Lavrov, M. (2005). Person identification from biological motion: Effects of structural and kinematic cues. *Perception & Psychophysics*, 67(4), 667–675. <http://dx.doi.org/10.3758/BF03193523>.
- Ullman, S. (1979). The interpretation of structure from motion. *Proceedings of the Royal Society B: Biological Sciences*, 203(1153), 405–426. <http://dx.doi.org/10.1098/rspb.1979.0006>.
- 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*. <http://dx.doi.org/10.3389/fpsyg.2014.00633>. 5(JUN).