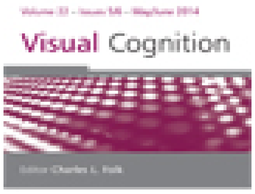

Visual Cognition

ISSN: 1350-6285 (Print) 1464-0716 (Online) Journal homepage: http://www.tandfonline.com/loi/pvis20

The role of familiarization in dynamic person recognition

Noa Simhi & Galit Yovel

To cite this article: Noa Simhi & Galit Yovel (2017): The role of familiarization in dynamic person recognition, Visual Cognition, DOI: 10.1080/13506285.2017.1307298

To link to this article: http://dx.doi.org/10.1080/13506285.2017.1307298

Published online: 10 Apr 2017.

VISUAL COGNITION, 2017
http://dx.doi.org/10.1080/13506285.2017.1307298

The role of familiarization in dynamic person recognition

Noa Simhi[a] and Galit Yovel[a,b]

[a]The School of Psychological Sciences, Tel Aviv University, Tel-Aviv, Israel; [b]The Sagol School of Neuroscience, Tel Aviv University, Tel-Aviv, Israel

ABSTRACT
Person recognition in real life is typically performed on dynamic whole people. Recent studies have indicated that the body and motion may contribute to person recognition beyond the face. In the current study, we examined the role of dynamic identity signatures—the idiosyncratic dynamic pattern of an individual—in unfamiliar and familiarized person recognition. To familiarize participants with the studied identities, we used an individuation training procedure and tested person recognition before and after training. In Experiment 1, participants studied people in motion and recognized them from dynamic or multi-static images (i.e., multiple still images taken from the video). We found that person recognition was better when people were recognized from dynamic than multi-static stimuli but only for familiarized people. These findings suggest that dynamic identity signatures may be used for familiar person recognition. In Experiment 2, we ruled out two alternative explanations by presenting multi-static images in study and testing recognition from dynamic or multi-static stimuli: lower recognition rates from dynamic stimuli following exposure to multi-static than dynamic stimuli indicated no evidence for a general advantage to recognition from video; lower recognition rates from multi-static stimuli following exposure to multi-static stimuli indicated no evidence for an advantage due to the congruence between the study and test media. We conclude that dynamic identity signatures may contribute to person recognition, but only of familiar people who were previously seen in motion.

ARTICLE HISTORY
Received 19 July 2016
Accepted 6 March 2017

KEYWORDS
Person recognition; face recognition; dynamic identity signatures; familiar faces; individuation; motion

In real life we typically recognize dynamic whole people. Nevertheless, the majority of person recognition studies have examined recognition of static faces. Thus, the question of whether person recognition relies also on the body and body motion has so far been scarcely investigated. Early studies that did explore person recognition from the whole person often presented videos of people in motion but tested recognition on static images of faces (Bruce, Henderson, Newman, & Burton, 2001; Burton, Wilson, Cowan, & Bruce, 1999; Davis & Valentine, 2009; Liu, Seetzen, Burton, & Chaudhuri, 2003; Pilz, Vuong, Bülthoff, & Thornton, 2011; Roark, O'Toole, Abdi, & Barrett, 2006; Schiff, Banka, & de Bordes Galdi, 1986). For example, in the first study that examined recognition of whole people from videos, Burton et al. (1999) presented subjects with videos of 20 different people in motion and asked subjects to recognize them from static high quality images of their faces. Results showed very poor performance when subjects where unfamiliar with the people presented in the videos. Using the same stimuli, Liu et al. (2003) presented high or low resolution faces at test and revealed little advantage to recognition from high resolution relative to low resolution faces. Bruce et al. (2001) also examined recognition of static images of faces following the presentation of videos or static images taken from the videos of the whole person and found poor recognition for unfamiliar faces and no advantage following presentation of people in videos relative to static images. Thus, recognition of static images of unfamiliar faces does not benefit from prior exposure to dynamic vs. static presentations of the whole person.

Later studies on whole person recognition examined whether the body and motion contribute to person recognition beyond the face by studying recognition from the whole person in comparison to the face alone. These studies have shown that the relative reliance on the face and body for whole person recognition varies as a function of two important factors: (1) whether the studied people are presented in dynamic or static images and (2) the degree of uncertainty in face recognition (for a review see

CONTACT Noa Simhi simhinoa@gmail.com School of Psychological Sciences & Sagol School of Neuroscience, Tel Aviv University, Tel Aviv 69978, Israel

Yovel and O'Toole, 2016). With respect to the role of dynamic information, O'Toole et al. (2011) have shown similar performance for face and whole person recognition from static images but better recognition from the whole person than the face with dynamic stimuli. These findings suggest that the body contributes to person recognition beyond the face for the dynamic but not static whole person. Simhi and Yovel (2016) similarly revealed that the body contributes to person recognition beyond the face following exposure to dynamic but not static stimuli also when performing recognition on static images.

In addition to the availability of dynamic information, the second factor that determines the relative reliance on the face and body in person recognition is the degree of uncertainty in face recognition. For example, the distance between the person and the camera has been shown to influence whether the body is used in person recognition: in a detailed investigation of the role of distance and exposure in dynamic person recognition, Hahn, O'Toole, and Phillips (2015) have shown that the body contributes to whole person recognition beyond the face—that is recognition was better for the whole person than the face alone—at larger distances but not at moderate or close distances, where recognition from the face and the whole person were comparable. Furthermore, examination of person recognition based on different segments from the videos indicated that participants primarily rely on the final image they see rather than accumulate information over time. Thus, when a close-up of the face is available, information from the body is not used at all. Another case in which uncertainty about the identity of the face led to greater reliance on information from the body was reported by Rice, Phillips, Natu, An, and O'Toole (2013) who demonstrated that, when face discrimination is difficult, the body is used for person recognition even in static images. Interestingly, participants were unaware that they were using the body rather than the face in these cases. Taken together, recognition of the whole person is likely to rely on both the face and the body, rather than primarily on the face, when people are seen in motion and when information from the face is less available.

One factor that has been suggested to mediate the role of motion in person recognition is *dynamic identity signatures* (Lander & Chuang, 2005; O'Toole & Roark, 2010; O'Toole, Roark, & Abdi, 2002). Dynamic identity signatures refer to idiosyncratic motion patterns of each individual, that can be learned and used in addition to any static information which may be available. In a recent study, Simhi and Yovel (2016) showed that dynamic identity signatures did not contribute to whole person recognition in an unfamiliar person matching task: person recognition from static images did not differ from recognition from a video following presentation of people in motion. These findings are consistent with Robbins and Coltheart (2015) who also used a similar task and did not find a difference in performance for recognition from dynamic and static unfamiliar whole person stimuli. Nevertheless, dynamic identity signatures may still contribute to the recognition of familiar people. Studies with famous faces have consistently shown that when face recognition is difficult, recognition from video is preferable to recognition from static images (Knight & Johnston, 1997; Lander & Bruce, 2000; Lander, Bruce, & Hill, 2001; Lander, Christie, & Bruce, 1999). Additionally, recognition of familiarized moving faces has also been shown to be preferable to recognition from static faces (Lander, Davies, Lander, & Davies, 2007). Recently, it has also been demonstrated that facial motion is more important for person recognition as familiarity with a face increases (Butcher & Lander, 2016).

With respect to the role of dynamic identity signatures in familiar whole person recognition, Pilz and Thornton (in press) showed that dynamic identity signatures were used for recognition of familiarized dynamic avatars in which facial identity was ambiguous and the body was uninformative. These findings indicate that dynamic identity signatures may be used when other information is not available. The advantage of motion for familiarized person recognition was also shown by Robbins and Coltheart (2015) who compared recognition of static and dynamic familiarized people in a naming task and found better recognition from dynamic than static whole person stimuli. However, two possible alternative explanations may account for this finding. First, it is possible that there is a general advantage to recognition of familiarized people from video. Such a possibility can be tested by examining recognition from video following exposure to multi-static images. Another possibility is a same-media recognition advantage, that is the study of dynamic people and recognition of people in video produces a recognition advantage since there is no need for between-media generalization. To examine this, it is necessary to test recognition of people from static images following the study of static images. In addition to these alternative explanations, it is important to note that Robbins and Coltheart (2015) used a naming task in order to assess familiarized person recognition. Several studies have shown that face recognition may be dissociated from naming, as participants may recognize previously seen individuals even if they cannot name them (e.g., Bruce & Young, 1986; Burton, Bruce, & Johnston, 1990; Schwartz & Yovel, 2016; Semenza, Zettin, & Borgo, 1998). Finally, a naming task cannot be used to assess unfamiliar person recognition and so Robbins and Coltheart (2015) could not compare unfamiliar and familiar person recognition using the same stimuli and design. In the current study, we therefore examined the possible contribution of dynamic identity signatures to familiarized person recognition using an old/new person recognition task which did not depend on naming (Experiment 1), and addressed the alternative explanations as well (Experiment 2).

To this end, we used a person recognition task that can be performed before and after a familiarization procedure in order to test the extent to which dynamic identity signatures may be used in the recognition of familiarized identities. Furthermore, to maximize reliance on the whole person rather than the face, we used videos of approaching walking people that did not include the face close-up, but ended at the point where the participant's feet were no longer visible (see example frames in Figure 1). In order to familiarize participants with the identities in the experiment, we presented participants with an individuation training procedure that has been shown to be effective for recognition of faces (McGugin, Tanaka, Lebrecht, Tarr, & Gauthier, 2011; Schwartz & Yovel, 2016; Tanaka & Pierce, 2009; Yovel et al., 2012), and examined recognition of the studied identities using an old/new task that was performed before and after the familiarization procedure. We designed a relatively challenging person recognition task so the relative effect of familiarization on performance across the different conditions would not be masked by a ceiling effect. The old/new task does not require naming and therefore examines person recognition independently of whether subjects may recall the person's name. Importantly, we assessed person recognition from new images of the same person to ensure that performance reflects the processing of person identity rather than image-specific matching. To assess whether dynamic information contributes to familiar person recognition beyond static information, in Experiment 1, two groups of participants were presented with dynamic people and were asked to recognize them before and after familiarization with the studied dynamic stimuli, from either dynamic or static images. In order to determine if the advantage to recognition of familiarized people from video results from a general advantage to familiarized person recognition from video or from same-media effects, in Experiment 2 we examine person recognition from videos and multi-static images after study of multi-static images alone.

Experiment 1

A contribution of dynamic identity signatures to person recognition would entail that, after studying a person in motion, recognizing that person in

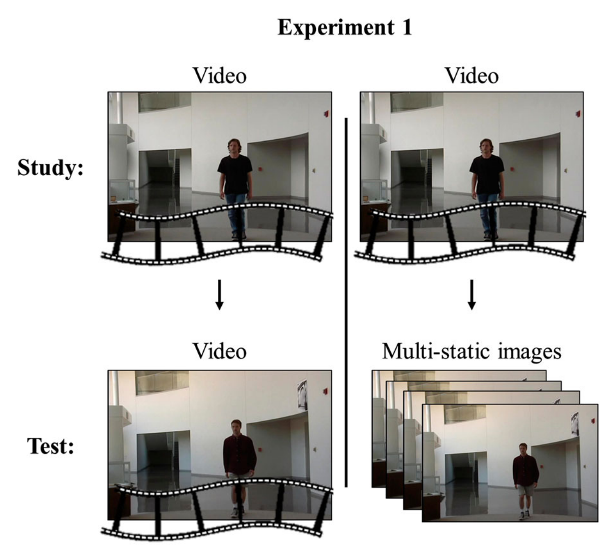

Figure 1. Schematic presentation of the design in Experiment 1. The study stimuli presented in this figure (top row) were presented both during the study and familiarization sessions of the experiment. The recognition stimuli (bottom row) were presented only during the recognition stage of the old/new task. Note that study and recognition stimuli were always filmed on different days. This example figure depicts the same identity both in the study and test stimuli. The stimuli in this figure and all stimuli in the experiment were taken from the Database of Moving Faces and People (O'Toole et al., 2005).

motion is better than recognizing them in static images alone (see Figure 1). To test the hypothesis that dynamic identity signatures contribute to recognition of familiar but not unfamiliar people, we compared the difference between recognition from dynamic or static images of people that were studied in video before and after familiarization.

Methods

Participants

Forty participants were recruited at Tel Aviv University to take part in this experiment. They were randomly assigned to the two conditions: 20 participants recognized people in videos and 20 recognized people from still images (mean age = 23.92 years, $SD = 2.05$, 31 female). Participants took part in the experiment either for course credit or payment. All participants had normal or corrected-to-normal vision, spoke Hebrew as a native language, and gave their informed consent to participate in the study by signing the appropriate consent form approved by the Tel Aviv University ethics committee.

Stimuli

The stimuli in the experiment were adapted from the Database of Moving Faces and People (O'Toole et al., 2005). Thirty identities were selected (15 male), 20 of which were used as identities to be studied (10 male) and 10 as foil identities during the recognition phase. Only identities with over 5 s of video available during which the whole person was visible were included in the experiment. Of these identities, only Caucasian subjects were selected (to avoid other race effects during person recognition) and care was taken to include only identities of roughly similar age (early twenties to thirties) and similar weight, height and body build. For each of the identities in the experiment, videos depicting the identity walking towards the observer in a naturally lit corridor were used for stimuli creation. For the studied identities, two such videos were used (each of them filmed on a different occasion, up to six months apart). One video was used for study stimulus creation and one video was used for recognition stimulus creation. It is important to note that since the study and test videos were filmed on different occasions, considerable differences in appearance between the first and second video of each identity could occur, thus requiring generalization in person recognition during the task. For the foil identities, only one video was used in order to create the stimulus displayed during the recognition phase.

In order to create the stimuli in the experiment, the original videos from the Database of Moving Faces and People (O'Toole et al., 2005) were cut so that they were all 5 s in length for the study videos and 4 s in length for the videos used for recognition. All video segments ended at the point where the feet of the person in the video were no longer visible which resulted in video segments in which the final frame in the video depicted the people in the video at a similar distance in all cases even though the distance in the starting frame was slightly different between the identities due to differences in the walking pace (i.e., identities with a faster walking pace started slightly further away than identities with a slower walking pace in the first frame of the video).

To create the multi-static images used for recognition in the recognition from multi-static image condition, one static frame was selected for each second of the video, starting from the last frame in the video. In cases where the person in the video was mid-stride, the closest frame depicting the person in a neutral pose was selected in order to minimize implied motion cues in the still images. The selected images were presented in a scrambled but predefined order during the recognition phase in the experiment, with the last frame in the video always appearing last in the multi-static presentation as well in order to avoid any possible recency effects. Each image was displayed for 1 s with a 0.1 s inter-stimulus interval between the images in order to match the amount of information available in a video as closely as possible while avoiding apparent motion effects. See Figure 1 for a depiction of the conditions in this experiment.

Both videos and static images were displayed at roughly 18.3° by 12.3° visual angle during the experiment and took up 720 × 480 pixels on the screen.

Design

During the experiment, participants first performed an old/new person recognition task on identities which they were exposed to only once. Afterwards they underwent an individuation training procedure to become familiar with the identities in the experiment. Finally, after the individuation training participants repeated the old/new task once again to determine if and to what extent familiarization improved person recognition. See Figure 2 for a schematic presentation of the experimental design.

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 × 1080 screen resolution, in front of which the participants were seated at a comfortable distance of approximately 60 cm.

During the experiment, participants completed two identical old/new task sessions separated by a familiarization session. The old/new task sessions included two phases: a study phase and a recognition phase.

Study phase. During the study phase the participants viewed the videos of 20 identities consecutively, separated by a 0.75 s fixation interval. Each identity was presented only once in the study phase, which was immediately followed by the recognition phase.

Recognition phase. During the recognition phase the participants viewed 30 identities in video or multi-static displays, depending on the experimental condition, and were asked to determine whether each identity was displayed in the study phase. The identities in this phase included 10 novel identities and the 20 identities which were viewed in the study phase, however during the recognition phase the studied identities were presented in stimuli which were filmed on a different day than the stimuli in the study phase (as detailed in the Stimuli section), thus requiring participants to generalize between the appearance in the study and recognition stimuli when recognizing the studied identities. After each 4 s stimulus, participants performed their old/new judgments by pressing corresponding keys on their keyboard. Response was followed by a blank screen which was presented for 0.5 s followed by 1 s of a fixation presentation before the next stimulus appeared.

The old/new task was repeated twice during the experiment—before and after the familiarization. The familiarization session included four different tasks (similar in design to Yovel et al., 2012) in which the participants in the experiment learned to associate a name with each of the identities in the experiment, as detailed below. The names in the experiment were all four letters long and each began with a unique letter. The familiarization session was performed first on 10 out of the 20 identities presented in the study phase of the old/new task and then repeated for the remaining 10 identities (see Figure 2 for a representation of the familiarization session procedure). The familiarization session took about an hour to complete.

Phase 1. During this part of the familiarization session the participants passively viewed the same stimuli which were presented during the study phase of the old/new task, however during the familiarization phase a name was assigned to each identity and this name was presented above the stimulus when the identity was displayed. Each stimulus was repeated five times during the task and participants were asked to memorize both the identities and their names while viewing the stimuli.

Phase 2. During this part of the familiarization session participants viewed each of the 10 identities in the familiarization phase once, along with the

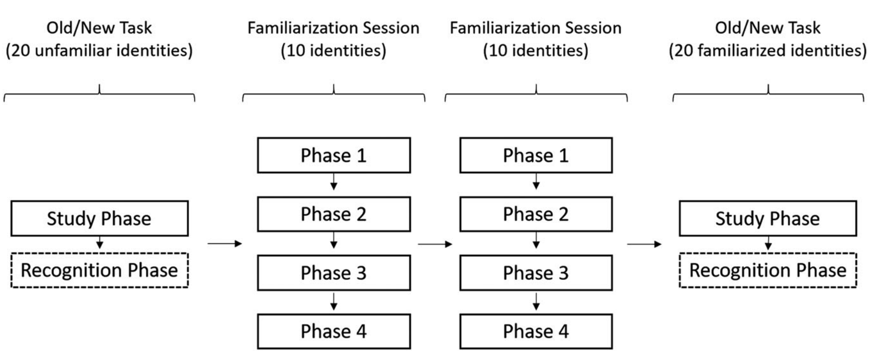

Figure 2. Schematic presentation of the experimental design. The experiment was divided into three main parts: two old/new tasks (each including a study and recognition phase) separated by a familiarization session (the familiarization session included four phases, as detailed in the Apparatus and procedure section, and was performed twice, each time on 10 different identities). Experimental phases marked by borders of the same style include stimuli filmed on the same day. The old stimuli presented in the recognition phase of the old/new task were of the same identities as those presented in the study phase, but from videos taken on a different day.

names assigned to them. Participants were required to press on their keyboard the first letter of the name of each identity that was presented when viewing them. If the correct letter was not pressed within the 5 s during which the video was displayed then the name remained on the screen until the correct key press.

Phase 3. Similar to Phase 2, during the third part of each familiarization session participants once again were required to press the letter on the keyboard which corresponded to the first letter of the name of the identity that was presented to them, however in Phase 3 the name was no longer presented on the screen and participants had to recall it from memory. If the wrong key was pressed, then participants were presented with the correct name on the screen and were required to press the correct key before the next stimulus appeared. Each stimulus was repeated six times during this task.

Phase 4. During the final phase of the individuation procedure participants were presented with a name which appeared in the centre of the screen for 1 s and was then replaced by a blank screen which appeared for 0.5 secs, followed by a video of one of the identities in the familiarization session. The task in this case was to indicate using "yes" and "no" keys whether the name which was presented matched the identity that followed it. During this task, each stimulus was repeated six times, and was randomly assigned to a match or mismatch trial during each repetition. Half of the trials in this task were match trials and half were mismatched names and identities.

The general design of the experiment is depicted in Figure 2. It is important to note that throughout the experiment two stimuli depicting each of the familiarized identities were used: one video which was used as a study stimulus and was presented during the old/new study phase and during the familiarization session, and one video or multi-static image set (depending on the experimental condition) which was used as the recognition stimulus and was presented only twice during the experiment, in the recognition phase of the old/new tasks, as can be seen in Figure 2.

Data analysis

Because the main purpose of our study was to assess the effect of familiarity on person recognition, our initial analysis focused on the proportion of correct recognition of the studied stimuli (Hits). We complemented this with an analysis of false alarm (FA) rates to evaluate the effect of motion on recognition of new images. We also examined criterion (C) and response times (RT) when correctly recognizing studied identities across the different conditions (RTs were measured from the time when the test stimulus disappeared until the time of response). Table 1 reports the mean and standard error of the mean of these measures as well as d' Additionally, the proportion of correct responses during the final familiarization phase, Phase 4, was analysed for each experimental condition to assess that participants indeed recognized the people following the familiarization session. Results show that the recognition level of the familiarized people reached ceiling performance (97–99% correct). Statistical analysis was performed using Statistica 9.0 (StatSoft Inc).

Results and discussion

To compare recognition rates for dynamic and static stimuli following exposure to video, before and after familiarization, we conducted a 2×2 mixed ANOVA on the proportion of hits, with test media (video vs. multi-static images) as a between subject factor and familiarity (unfamiliar vs. familiarized) as a within subject factor. This analysis revealed a main effect of test media type ($F(1,38) = 5.21$, $p = .03$, $\eta_p^2 = .12$), a main effect of familiarity ($F(1,38) = 65.21$, $p = 9.07*10^{-10}$, $\eta_p^2 = .63$) and an interaction ($F(1,38) = 6.72$, $p = .01$, $\eta_p^2 = .15$).

Post-hoc Tukey HSD analysis revealed that familiarization significantly improved person recognition in both the video ($p = .0002$) and multi-static ($p = .002$) conditions. Importantly, however, recognition from video was better than recognition from static images alone after familiarization ($p = .006$) but not before familiarization ($p = .99$), indicating that there is an advantage to recognizing people in motion only after familiarization (see Figure 3).

We further examined performance in this task using the proportion of FAs. The same 2×2 mixed ANOVA on FA rates reveals a lower FA rate for familiarized than unfamiliar identities ($F(1,38) = 24.99$, $p = 1.3*10^{-5}$, $\eta_p^2 = .4$). No effect of test media ($F(1,38) < 1$) or interaction ($F(1,38) < 1$) was found.

An analysis of the criterion, C, of participants in the task revealed no main effect of test media ($F(1,38) < 1$),

Table 1. The average and standard error of hits, false alarms (FA), *d′*, criterion (C) and response times (RT) for correctly recognized studied identities in Experiment 1 and Experiment 2.

	Experiment 1				Experiment 2			
Study	Video				Multi-static display			
Test	Video		Multi-static display		Video		Multi-static display	
	Unfamiliar	Familiarized	Unfamiliar	Familiarized	Unfamiliar	Familiarized	Unfamiliar	Familiarized
Hits	.55 (.03)	.82 (.02)	.54 (.03)	.68 (.03)	.55 (.02)	.73 (.02)	.52 (.02)	.73 (.02)
FA	.25 (.04)	.11 (.03)	.28 (.03)	.14 (.03)	.38 (.04)	.22 (.02)	.31 (.04)	.15 (.03)
d′	1 (.19)	2.4 (.15)	.78 (.13)	1.75 (.12)	.47 (.14)	1.52 (.13)	.64 (.16)	1.91 (.18)
C	.36 (.09)	.22 (.09)	.27 (.08)	.37 (.09)	.11 (.05)	.1 (.07)	.26 (.06)	.3 (.07)
RT	.8 (.05)	.77 (.06)	.82 (.08)	.74 (.12)	.87 (.08)	.76 (.05)	.69 (.05)	.7 (.012)

familiarity ($F(1,38) < 1$) or interaction ($F(1,38) = 2.4$, $p = .13$, $\eta_p^2 = .06$).

Analysis of RTs revealed no main effect of test media ($F(1,38) < 1$), familiarity ($F(1,38) < 1$) or interaction ($F(1,38) < 1$). See Table 1 for the average proportion of hits, FAs, *d′*, *C* and RTs in each condition.

Taken together these results show that for unfamiliar people there is no advantage to recognizing a person that was seen in motion from video as compared to static images, similar to the findings in the matching tasks in Robbins and Coltheart (2015) and Simhi and Yovel (2016). On the other hand, recognition of familiarized people that were studied in motion is better when they are recognized from dynamic than from multi-static images. This effect was limited to recognition of the studied stimuli; it was not found for correct rejection of new stimuli.

These findings suggest that people use dynamic identity signatures in familiarized person recognition. Nonetheless, two alternative explanations that may account for these findings must be ruled out. First, recognition of familiarized people from video may always be preferable to recognition from static images, since videos contain more information than static images alone. To examine this possibility, in Experiment 2 we presented subjects with multi-static images at study and tested their recognition with videos (see Figure 4, left). If the advantage to familiar person recognition from video is not due to the contribution of dynamic identity signatures, we would expect recognition in this condition to be comparable to recognition from video after the study of videos.

A second possibility is that recognition of people from videos after studying videos is preferable to recognition from static images because the study and

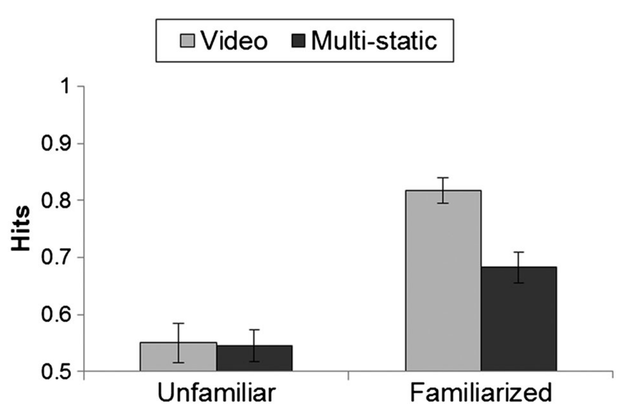

Figure 3. The proportion of correct recognition of studied identities (hits) in the old/new task for unfamiliar and familiarized people in Experiment 1. Error bars indicate the standard error of the mean.

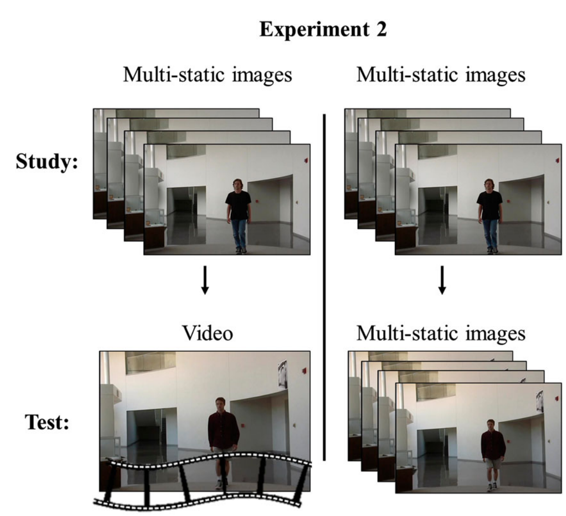

Figure 4. Schematic presentation of the conditions in Experiment 2. The study stimuli (top row) were presented both during the study and familiarization sessions of the experiment, and the recognition stimuli (bottom row) were presented only during the recognition stage of the old/new task in the experiment. Recognition from video after the study of multi-static images allowed us to rule out the possibility that there is a general advantage for person recognition from video. Recognition from multi-static images after the study of multi-static images allowed us to rule out the possibility that there is a same-media recognition advantage.

test stimuli are of the same media so there is no need for between-media generalization. To test this possibility, in Experiment 2 we also examined person recognition from multi-static images after the study of multi-static images (see Figure 4, right). If same-media recognition explains the advantage to recognition from video in Experiment 1, then we would expect recognition of familiarized people to be comparable regardless of whether they are studied and recognized from multi-static images or videos.

In addition to dynamic identity signatures, motion may contribute to person recognition through form-from-motion processes (known also as structure from motion) (see O'Toole & Roark, 2010; O'Toole et al., 2002; Yovel & O'Toole, 2016). Such processes have long been described in the context of object perception (see Koenderink, 1986; Ullman, 1979). In the context of faces they have been discussed as the representation enhancement hypothesis (O'Toole et al., 2002), postulating that for unfamiliar face recognition especially, exposure to the face in motion may improve the representation of that face and facilitate facial recognition also from static images. In the context of whole person recognition we have recently demonstrated that form-from-motion processes can indeed contribute to person recognition by improving the representation of the body and thereby facilitating whole person recognition in a matching task conducted on unfamiliar people (Simhi & Yovel, 2016). The conditions included in Experiments 1 and 2 also allowed us to examine the contribution of form-from-motion processes to unfamiliar and familiarized person recognition by comparing recognition from multi-static images following exposure to video or multi-static images at study. Better recognition from multi-static images following exposure to motion than following exposure to multi-static images would indicate a contribution of form-from-motion processes to person recognition.

Experiment 2

Methods

Participants

Forty participants were recruited at Tel Aviv University to take part in this experiment, either for course credit or payment. They were randomly assigned to one of two conditions: after the study of multi-still images 20 participants recognized people in videos and 20 recognized people from multi-still images (mean age = 23.57 years, $SD = 2.31$, 32 female). All participants had normal or corrected-to-normal vision, spoke Hebrew as a native language, and gave their informed consent to participate in the study by signing the appropriate consent form approved by the Tel Aviv University ethics committee.

Stimuli

In this experiment the stimuli in the study were multi-static images, similar to those that were used in the test of Experiment 1 in the recognition from multi-static image condition. In order to create these displays, four frames were extracted from the study videos, starting from the final frame in the video, at equally spaced intervals. During the experiment these images were presented in a scrambled but pre-defined order (with the final frame appearing last), for 1.25 s each with a 0.1 s interval.

Apart from this difference, the test stimuli in Experiment 2 were the same videos and multi-static displays used in Experiment 1 (see Figure 4 for a schematic presentation of the conditions in Experiment 2) and the design and procedure in Experiment 2 were identical to the previous experiment.

Data analysis

Data analysis in Experiment 2 was performed in the same manner as in Experiment 1. An analysis of the proportion of correct responses during the final familiarization phase, Phase 4, revealed that participants reached ceiling performance (98–99%).

Results and discussion

To assess whether there is a video advantage at test and/or a same-media effect in recognition of familiarized people, we performed a one-way ANOVA on the proportion of hits when recognizing familiarized people, with study-test media condition (Video–Video, Multi-static–Video, Multi-static–Multi-static) as a between participant factor. This analysis revealed a main effect of study-test media condition ($F(1,57) = 5.15$, $p = .009$, $\eta_p^2 = .15$). Post-hoc Tukey HSD analysis revealed that this effect was due to better recognition in the video–video condition than the multi-static–video condition ($p = .02$), and better recognition in the video–video condition than the multi-static–multi-static condition ($p = .02$). There was no

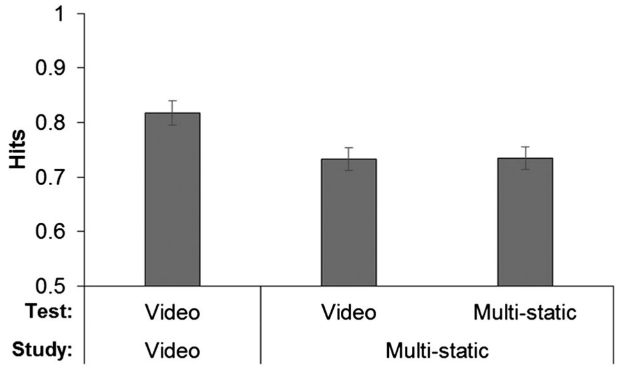

Figure 5. The proportion of correct recognition of studied identities (hits) when recognizing familiar people in the video–video condition in Experiment 1 and in the conditions in Experiment 2. Error bars indicate the standard error of the mean.

difference between the multi-static–video and multi-static–multi-static conditions ($p = .99$) (see Figure 5).

The advantage to recognition from video after the study of videos as compared to the study of multi-static images indicates that there is no overall advantage to recognition from video at test. Furthermore, the fact that recognition from video was not preferable to recognition from multi-static images alone also when multi-static images were studied in both cases, again indicates that the addition of dynamic information at test does not provide a general advantage in person recognition. Thus, these findings support the hypothesis that the advantage to recognition of familiarized people from video reflects the contribution of dynamic identity signatures to recognition. The finding that recognition from video after the study of videos was better than recognition from multi-static images after the study of multi-static images further indicates that the advantage to recognition from video in Experiment 1 did not result from same-media effects, as similar recognition could not be achieved from static images alone. The lack of a difference between recognition from video and multi-static images after the study of multi-static images supports this conclusion as well, again strengthening the evidence that a same-media advantage cannot explain these results.

A similar analysis on FAs also revealed a main effect of condition ($F(1,57) = 3.44$, $p = .04$, $\eta_p^2 = .11$). Post-hoc analysis revealed that this effect was due to significantly higher FAs for dynamic stimuli that were studied in multi-static images as compared to stimuli that were studied in video ($p = .03$). Thus, presentation of a smaller amount of information at study (multi-static images) as compared to test (videos) generated higher FAs in this condition. There was no difference between FAs in recognition from video and recognition from multi-static images when multi-static images were used at study (.23) indicating no general benefit for correctly rejecting new images presented in same media. Furthermore, there was no difference in FAs between recognition from video and multi-static stimuli when the same media was presented at study and test ($p = .61$). Thus, the main differences between these conditions were driven by differences in the recognition of studied identities and not the correct rejection of new identities.

An analysis of the criterion, *C*, of response in this task revealed no differences between the study-test media conditions ($F(1,57) = 1.65$, $p = .2$, $\eta_p^2 = .05$). Finally, no difference was found in RTs to correctly recognizing studied identities in the task ($F(1,57) < 1$). Table 1 summarizes the results of Hits, FA, *d′*, *C* and RT across the different conditions.

The conditions included in this experiment allowed us to also assess the role of form-from-motion processes by comparing recognition from multi-static images after the study of videos or multi-static stimuli. We conducted a 2 × 2 mixed ANOVA on the proportion of hits, with study media (video vs. multi-static images) as a between subject factor and familiarity (unfamiliar vs. familiarized) as a within subject factor in order to assess the possible contribution of form-from-motion processes to person recognition. This analysis revealed a main effect of familiarity ($F(1,38) = 53.97$, $p = 8.4*10^{-9}$, $\eta_p^2 = .59$), but no main effect of study media ($F(1,38) < 1$) or interaction ($F(1,38) = 2.61$, $p = .11$, $\eta_p^2 = .06$). These findings suggest that recognition from static images was not better after studying dynamic than static people in both unfamiliar and familiarized person recognition.

An analysis of FAs revealed a main effect of familiarity ($F(1,38) = 23.88$, $p = 1.9*10^{-5}$, $\eta_p^2 = .39$), but no main effect of study media ($F(1,38) < 1$) or interaction ($F(1,38) < 1$). An analysis of criterion revealed no main effects of familiarity ($F(1,38) = 1.08$, $p = .3$, $\eta_p^2 = .03$), study media ($F(1,38) < 1$) or interaction ($F(1,38) < 1$). Finally, an analysis of RTs also revealed no main effects of familiarity ($F(1,38) < 1$), study media ($F(1,38) < 1$) or interaction ($F(1,38) < 1$). These findings indicate that form-from-motion processes did not contribute to unfamiliar and familiarized person recognition. We discuss these findings in light of previous studies that did reveal the contribution of form-from-motion processes to unfamiliar person recognition in a sequential matching task (Simhi & Yovel, 2016) in the General discussion.

Finally, we included all conditions presented in Experiments 1 and 2 in a 2 × 4 mixed ANOVA with familiarity (unfamiliar vs. familiarized) as a within subject factor and condition (Video–Video, Video–Multi-static, Multi-static–Video and Multi-static–Multi-static) as a between subject factor with hit rate as the dependent measure. This analysis revealed a main effect of familiarity ($F(1,76) = 151.29$, $p < 10^{-16}$, $\eta_p^2 = .67$) and an interaction between familiarity and condition ($F(1,76) = 2.82$, $p = .04$, $\eta_p^2 = .1$). A separate analysis of hits before and after familiarization showed a significant difference in performance across the four conditions after familiarization ($F(1,76) = 6.06$, $p = .0009$, $\eta_p^2 = .19$) but not before familiarization ($F(1,76) < 1$). These effects are consistent with the results reported above, showing better recognition of familiarized but not unfamiliar dynamic people following exposure to dynamic than static images.

General discussion

The goal of this study was to assess the role of motion in unfamiliar and familiarized person recognition. We found that, after studying people in motion, recognition was better when people were presented in dynamic than in multi-static images indicating that dynamic identity signatures contribute to familiar person recognition. This benefit of recognizing people in motion was found for familiarized people, but not for unfamiliar people. Notably, we ruled out two alternative explanations for these findings. First, we demonstrated that the benefit of recognizing dynamic people was not due to a general advantage of recognition of familiarized people from motion, as the recognition rate of people from motion was lower when they were studied from multi-static images as compared to when they were studied from dynamic stimuli. Second, we demonstrated that the advantage to video recognition did not result from a same-media recognition advantage as studying and recognizing people in multi-static images did not improve person recognition at a similar rate as when people were studied and recognized in video, again indicating that the advantage to recognition of dynamic people resulted from the use of the unique motion cues in the video itself.

While previous studies have demonstrated that dynamic identity signatures do not contribute to unfamiliar person recognition (Robbins & Coltheart, 2015; Simhi & Yovel, 2016), to our knowledge this is the first study that examined unfamiliar and familiarized whole person recognition within the same design. Our study therefore provides compelling evidence that dynamic identity signatures are studied over time and contribute to whole person recognition when we are familiar with the person being recognized (see O'Toole & Roark, 2010; O'Toole et al., 2002; Yovel & O'Toole, 2016).

It is important to note that additional factors, other than familiarity, may also influence whether dynamic identity signatures are used in person recognition. One such factor is the quality of the information being presented. Studies with dynamic faces show that the contribution of motion to familiar person recognition is usually found for degraded videos (e.g., Butcher & Lander, 2016; Knight & Johnston, 1997; Lander & Bruce, 2000; Lander et al., 1999, 2001). In the context of the whole person, it has recently been shown that when familiarized avatars with ambiguous facial identity and uninformative bodies are presented then dynamic identity signatures are used for person recognition (Pilz & Thornton, in press). Finally, studies with point light displays (see Johansson, 1973) also indicate that participants can recognize familiar people based on their body movement pattern (e.g., Cutting & Kozlowski, 1977; Hill & Pollick, 2000; Jacobs, Pinto, & Shiffrar, 2004; Loula, Prasad, Harber, & Shiffrar, 2005; Troje, Westhoff, & Lavrov, 2005) using the point light information alone. Taken together, these studies suggest that dynamic identity signatures may be used for recognition especially under impoverished conditions when information from other cues such as the face is less available. For example, recognition of people in low illuminations or for people with low vision is more likely to depend on general movement patterns rather than when high resolution face and body information is available. It should also be noted that recognition of people from point light displays was better from videos of jumping and dancing rather than for walking stimuli (Loula et al., 2005). Such movements are less common in real life but are more variable in terms of the dynamic information they convey and therefore are more likely to convey clear dynamic identity signatures that may provide a greater motion advantage in person recognition.

Notably, dynamic identity signatures (O'Toole & Roark, 2010; O'Toole et al., 2002) are not the only possible way through which motion can contribute to person recognition. Dynamic mental representations (Freyd, 1987) outline the importance of motion even in the perception of static images which contain implicit dynamic information. Form-from-motion, or structure from motion processes, may also contribute to person recognition of previously seen people in motion. Indeed, we have recently shown that recognition of unfamiliar people from static images in a sequential matching task was better following exposure to dynamic than static images (Simhi & Yovel, 2016). This advantage was not found in the current study before familiarization. There are several possible explanations that may account for this inconsistency. First, person recognition before familiarization was relatively low, and therefore it is possible that the task was too difficult to reveal a benefit of motion. Second, the benefit of motion may not emerge in an old–new task, in which the delay between the presentation of study and test images is much longer. It is important to note that the same explanations may not apply to the lack of benefit from dynamic identity signatures for unfamiliar person recognition, as this finding is consistent with previous studies which used a sequential matching task (Robbins & Coltheart, 2015; Simhi & Yovel, 2016) and yielded a relatively high performance level. While Simhi and Yovel (2016) did reveal a benefit of form-from-motion processes to person recognition, no contribution of dynamic identity signatures in unfamiliar person recognition was found in either Simhi and Yovel (2016) or Robbins and Coltheart (2015). Future studies may further examine the conditions required to reveal a benefit of form-from-motion processes in person recognition.

To familiarize subjects with the stimuli, we applied an individuation training procedure that has been shown to be effective in previous experiments using face tasks in generating an invariant representation that supports recognition of faces from different views or lightings that were not presented during the study phase (McGugin et al., 2011; Schwartz & Yovel, 2016; Tanaka & Pierce, 2009; Yovel et al., 2012). Here we applied this training to a relatively challenging whole person recognition task that was created by selecting stimuli depicting identities that were within the same age range and had relatively similar weight and height and yielded relatively low recognition rates before familiarization. Interestingly, following individuation training in which each of the 20 identities that were presented in the study was associated with a different name, recognition of the studied identities increased by 20–35%. It should be noted that the familiarization included no more than 20 repetitions of each of the different identities across the four phases of the training. Thus, the results of our study also reveal that relatively limited exposure to each person in which each identity is associated with a name had a significant effect on recognition of these people from new stimuli that were significantly different from those that were studied and were hardly recognized before training.

Our findings may also have implications for suspect recognition, which often involves matching closed circuit television videos (CCTV) or eyewitness memory for real-life crime scenes to static images of suspects. For unfamiliar person recognition, our data and the findings of Robbins and Coltheart (2015) and Simhi and Yovel (2016) suggest that, after exposure to a person in motion, recognition from dynamic or static displays can be equally effective. In regards to familiar person recognition, the preferable media may depend on previous exposure. To give an example, if a jury is exposed to many repetitions of a documented crime during a trial, these results suggest that if the crime was documented in video they may be more accurate in determining if a suspect was present at the crime scene when doing so from a video, where dynamic identity signatures may assist them in determining the suspect's identity; if only photographs exist from the crime scene however then recognition from static images and videos may be equally accurate. Given that we are usually exposed to the dynamic whole person, in real life situations recognition of familiar people from video may usually be preferable to recognition from static images.

In summary, our findings show that dynamic identity signatures are used in familiar whole person recognition, and these idiosyncratic motion cues are most likely studied over time. Furthermore, we show that relatively short individuation training with dynamic and static images of the whole person can significantly improve person recognition in challenging person recognition tasks. The dynamic whole person is now only beginning to be studied, and the factors which influence how motion effects person perception should continue to be explored.

Acknowledgements

We would like to thank Alice O'Toole for providing us with the video database of moving faces and people. Special thanks also to undergraduate students Itai Linzen, Gili Katabi, Gal Safra, Guy Brillant and Talya Rapp for their assistance in data collection.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Israel Science Foundation (ISF) under Grant 392/16 to GY.

References

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*(4), 433–436. doi: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. doi:10.1037//1076-898X.7.3.207

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology (London, England : 1953)*, *77*(Pt 3), 305–327. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3756376

Burton, A. M., Bruce, V., & Johnston, R. A. (1990). Understanding face recognition with an interactive activation model. *British Journal of Psychology (London, England : 1953)*, *81*(Pt 3), 361–380. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2224396

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. doi:10.1111/1467-9280.00144

Butcher, N., & Lander, K. (2016). Exploring the motion advantage: Evaluating the contribution of familiarity and differences in facial motion. *The Quarterly Journal of Experimental Psychology*, *218*(February), 1–30. doi:10.1080/17470218.2016.1138974

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. doi:10.3758/BF03337021

Davis, J. P., & Valentine, T. (2009). CCTV on trial: Matching video images with the defendant in the dock. *Applied Cognitive Psychology*, *23*(4), 482–505. doi:10.1002/acp.1490

Freyd, J. J. (1987). Dynamic mental representations. *Psychological Review*, *94*(4), 427–438. doi: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 (London, England : 1953)*, 1–18. doi:10.1111/bjop.12125

Hill, H., & Pollick, F. E. (2000). Exaggerating temporal differences enhances recognition of individuals from point light displays. *Psychological Science*, *11*(3), 223–228. doi: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. doi: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. doi:10.3758/BF03212378

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in psychtoolbox-3. *Perception*, *36*(14), 1–16.

Knight, B., & Johnston, A. (1997). The role of movement in face recognition. *Visual Cognition*, *4*(3), 265–273. doi:10.1080/713756764

Koenderink, J. J. (1986). Optic flow. *Vision Research*, *26*(1), 161–179. doi:10.1016/0042-6989(86)90078-7

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., & Hill, H. (2001). Evaluating the effectiveness of pixelation and blurring on masking the identity of familiar faces. *Applied Cognitive Psychology*, *15*(1), 101–116. doi:10.1002/1099-0720(200101/02)15:1<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 & Cognition*, *27*(6), 974–985. doi: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

Lander, K., Davies, R., Lander, K., & Davies, R. (2007). Exploring the role of characteristic motion when learning new faces. *The Quarterly Journal of Experimental Psychology*, *60*(4), 519–526. doi:10.1080/17470210601117559

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. doi: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. doi:10.1037/0096-1523.31.1.210

McGugin, R. W., Tanaka, J. W., Lebrecht, S., Tarr, M. J., & Gauthier, I. (2011). Race-Specific perceptual discrimination improvement following short individuation training with faces. *Cognitive Science*, *35*(2), 330–347. doi:10.1111/j.1551-6709.2010.01148.x

O'Toole, A. J., Harms, J., Snow, S. L., Hurst, D. R., Pappas, M. R., Ayyad, J. H., & Abdi, H. (2005). A video database of moving faces and people. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *27*(5), 812–816. doi:10.1109/TPAMI.2005.90

O'Toole, A. J., Phillips, P. J., Weimer, S., Roark, D. A., Ayyad, J., Barwick, R., & Dunlop, J. (2011). Recognizing people from dynamic and static faces and bodies: Dissecting identity with a fusion approach. *Vision Research*, *51*(1), 74–83. doi:10.1016/j.visres.2010.09.035

O'Toole, A. J., & Roark, D. A. (2010). Memory for moving faces: The interplay of Two recognition systems. In *Dynamic faces*. The MIT Press. doi:10.7551/mitpress/9780262014533.003.0003

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. doi:10.1016/S1364-6613(02)01908-3

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442. doi:10.1163/156856897X00366

Pilz, K. S., & Thornton, I. M. (in press). Idiosyncratic body motion influences person recognition. *Visual Cognition*. doi:10.1080/13506285.2016.1232327

Pilz, K. S., Vuong, Q. C., Bülthoff, H. H., & Thornton, I. M. (2011). Walk this way: Approaching bodies can influence the processing of faces. *Cognition*, *118*(1), 17–31. doi: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. doi:10.1177/0956797613492986

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. doi:10.1068/p5503

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. doi:10.1016/j.jecp.2015.04.006

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), 219–231. doi:10.2307/1422276

Schwartz, L., & Yovel, G. (2016). The roles of perceptual and conceptual information in face recognition. *Journal of Experimental Psychology: General*, *145*(11), 1493–1511. doi:10.1037/xge0000220

Semenza, C., Zettin, M., & Borgo, F. (1998). Names and identification: An access problem. *Neurocase*, *4*(1), 45–53. doi:10.1080/13554799808410606

Simhi, N., & Yovel, G. (2016). The contribution of the body and motion to whole person recognition. *Vision Research*, *122*, 12–20. doi:10.1016/j.visres.2016.02.003

Tanaka, J. W., & Pierce, L. J. (2009). The neural plasticity of other-race face recognition. *Cognitive, Affective, & Behavioral Neuroscience*, *9*(1), 122–131. doi:10.3758/CABN.9.1.122

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. doi:10.3758/BF03193523

Ullman, S. (1979). The interpretation of structure from motion. *Proceedings of the Royal Society B: Biological Sciences*, *203*(1153), 405–426. doi:10.1098/rspb.1979.0006

Yovel, G., Halsband, K., Pelleg, M., Farkash, N., Gal, B., & Goshen-Gottstein, Y. (2012). Can massive but passive exposure to faces contribute to face recognition abilities? *Journal of Experimental Psychology: Human Perception and Performance*, *38*(2), 285–289. doi:10.1037/a0027077

Yovel, G., & O'Toole, A. J. (2016). Recognizing people in motion. *Trends in Cognitive Sciences*, *20*(5), 383–395. doi:10.1016/j.tics.2016.02.005