A search advantage for dynamic same- and other-race faces

Natalie Butcher[1], Karen Lander[2] & Rachel Jagger[2]

[1]*Social Futures Institute, Teesside University, Middlesbrough, United Kingdom, TS1 3BX*

[2] *School of Psychological Sciences, University of Manchester, Manchester, United Kingdom,*

M13 9PL

WORD COUNT (excluding abstract, title page, references & figures): 7919

Correspondence concerning this article should be addressed to: Dr Natalie Butcher,

n.butcher@tees.ac.uk

Abstract

Previous work has found faster and more accurate recognition of faces learnt from moving sequences than from static images. However, it is unclear whether *the motion advantage* can be generalised to other-race face learning. In two experiments, we examined the motion advantage for same- and other-race faces in the context of a delayed visual search task and investigated the importance of face repetitions and clip length at learning. Participants learnt faces as a static image or a moving clip and then searched for these target faces in visual arrays. Same-race target faces were located faster and more accurately than other-race targets. A dynamic search advantage was revealed for both same- and other-race faces, whereby search latencies were shorter, and accuracy higher for faces learnt in motion. Furthermore, differing clip lengths and face repetitions during familiarisation yielded the same dynamic advantage (Experiment 2), suggesting that motion provides a robust and valuable cue for the identification of both same- and other-race faces.

Keywords: face recognition, biological motion, visual search, race

Human face recognition is not perfect and struggles to perform under certain conditions (Braje Kersten, Tarr, & Troje, 1998; Hill, Schyns, Akamatsu, 1997; Troje & Bülthoff, 1996; Wagenaar & van der Schrier, 1996). For instance recognition is more difficult for previously unfamiliar faces than familiar faces (Henderson, Bruce, & Burton, 2001; Pike, Kemp, Towell & Philips, 1997) and is particularly poor for other-race rather than same-race faces ('other-race effect', see Meissner & Brigham, 2001).

Facial motion has also been found to affect face processing with movement facilitating the recognition of both highly familiar and previously unknown faces ('motion advantage', see Hill & Johnston, 2001; Knappmeyer, Thornton & Bülthoff, 2003; Knight & Johnston, 1997; Lander, Christie & Bruce, 1999; O'Toole, Roark & Abdi, 2002; Thornton & Kourtzi, 2002). Faces perform two types of movement; a) non-rigid, transient structural transformations of the facial skeletal musculature, and b) rigid, transient changes in face orientation while the facial structure remains unchanged. These head and facial movements provide valuable cues to facial identification that are not available from static images, including 3D structural information from motion, and dynamic motion signatures (see O'Toole et al., 2002 for theoretical accounts of motion advantage). Although the beneficial effects of motion are robust for familiar face recognition (Knight & Johnston, 1997; Lander et al., 1999), the effect of motion during learning and recognition of unfamiliar faces is less consistent (see Lander & Bruce, 2003, Pike et al., 1997; Thornton & Kourtzi, 2002; who reported a motion advantage and; Christie & Bruce, 1998 who reported no motion advantage).

Inconsistencies within the unfamiliar face literature regarding the motion advantage have led several researchers to consider why contradictory results arise. One possibility is the nature of the experimental task being utilised (Pilz, Thornton & Bülthoff, 2006). The standard paradigm used to investigate the effect of facial motion on unfamiliar face recognition is the old/ new recognition memory task. Pilz et al. (2006, p.436) argue that this paradigm may make

it "difficult to extract stable measures of dynamic performance" as it biases observers to assume techniques dependent on memorization that favour static content. To overcome this bias and place more emphasis on working memory representations, Thornton and Kourtzi (2002) used an immediate matching paradigm to demonstrate the short-term differential impact of priming with a short movie clip compared to a static image. Responses following a dynamic prime were found to be faster than those following static primes. Pilz et al. (2006) also used a novel approach to investigate the motion advantage whilst avoiding possible drawbacks associated with old/ new recognitions tasks. Using a delayed visual search paradigm they explored how learning over a longer period of time is affected by the availability of non-rigid motion. Participants were familiarised with two target faces; one static and one moving. Participants then completed a visual search task in which they were required to indicate whether either of the target faces was present. Pilz et al. (2006) found a reliable dynamic search advantage such that observers identified faces learnt from dynamic sequences faster than those learnt in static (also see Pilz, Bülthoff & Vuong, 2009). Thus, a motion advantage for unfamiliar faces has been demonstrated across a number of experimental paradigms (old/new recognition; matching tasks and visual search) and impacts on both short-term and long-term identity decisions. It is hypothesised that when learning previously unfamiliar faces, motion facilitates the construction of more robust mental representations at encoding which later aid recognition; the 'representation enhancement hypothesis' (O'Toole et al., 2002).

One shortcoming of the current literature on the motion advantage is that existing studies have almost exclusively used same-race facial stimuli despite evidence that race affects face recognition (for review see; Hugenberg, Young, Bernstein & Sacco, 2010). Therefore, it remains unclear whether the motion advantage can be generalised to other-race face learning. Similarly, existing studies of the other race effect in face recognition have, until recently (Zhao, Hayward & Bülthoff, 2014), used single static images, highlighting the importance of studying

the other-race effect in relation to facial motion. Although the other race effect in face recognition is a widely researched phenomenon, confirmed by several meta-analytic studies (Bothwell, Brigham & Malpass, 1989; Meissner & Brigham, 2001) very little is known about whether facial motion acts as a cue to identity recognition for other-race faces as for same-race faces. So, it is important to consider the mechanisms through which facial motion might affect other-race face recognition in particular. That is not to say that well-documented explanations of the motion advantage (e.g. motion at encoding leading to more robust face representations) are not likely to play a role in the effect of motion on other-race face recognition, but that current literature offer insight to particular reasons why motion may / may not benefit the recognition of other-race faces. We will now consider such mechanisms.

First, it is possible that other-race observers lack the perceptual expertise to process and utilise other-race facial motion to aid recognition. There is no universally accepted agreement of the mechanisms responsible for the other race effect (Meissner, Brigham, & Butz, 2005) but it has been posited that other-race observers may lack the perceptual expertise to process these faces. Such claims are based on findings that other-race face processing lacks the hallmarks of same-race face processing; a) holistic face processing (i.e. perceiving the face as a whole rather than a collection of independent face parts) and b) processing of configural information (i.e. spacing between face features) (e.g. Harrison et al., 2011; Johnson & Fredrickson, 2005; Rhodes, Tan, Brake & Taylor, 1989; Tanaka, Kiefer & Bukach, 2004; but see Hayward, Crookes, & Rhodes, 2013). For same-race faces, several empirical findings are associated with holistic processing; face inversion (Yin, 1969), the composite effect (Young, Hellawell & Hay, 1987) and the peripheral identification technique (McKone 2004). Studies using these tasks suggest that holistic processing for other-race faces is weaker than for same-race faces (e.g. Tanaka et al., 2004), although results are somewhat mixed (see Valentine & Bruce, 1986). It is apparent that differential experience with same-race and other-race faces leads to qualitatively

different processing styles, differences that may develop during infancy (Wheeler et al., 2011). Using the face composite effect Xiao and colleagues (Xiao, Quinn, Ge & Lee, 2012, 2013) investigated what aspect of face recognition is influenced by facial motion information; featural (i.e. processing individual face parts) or holistic processing. Their findings suggest that dynamic facial information (rigid and non-rigid) leads to face processing in a featural or part-based manner. They found a standard face composite effect in their static conditions but no face composite effect in their dynamic conditions. Therefore, it is possible that facial movement may enhance the reliance on featural processing for other-race faces. However, the tasks implemented by Xiao et al., (2012, 2013) required participants to identify the top half of a composite face wherein the bottom half always belonged to a foil face, rather than identifying a whole face. Thus, the extent to which these findings can be generalized to other face recognition tasks (that don't favour part-based processing) is unclear. Taken together, if observers lack the perceptual expertise to process other-race faces and facial movement compounds these processing differences then other-race observers are unlikely to benefit from facial motion and thus, we should not observe a motion advantage for other-race faces.

Alternatively, there is some evidence to suggest that facial motion may facilitate holistic and configural processing more specifically (Thornton, Mullins & Banahan, 2011; Xiao, et al., 2015). Xiao et al., (2015) found facial motion promoted more fixation shifting (in 9 month infants) across the whole face region and that scanning behaviour was related to recognition performance. It is argued that facial movement alters the face configuration from moment to moment, offering the observer a range of possible facial configurations that is not gained from a static image. By processing these possible facial configurations the observer might be better prepared to recognise that face. As such the availability of facial motion may facilitate, the arguably disrupted, configural processing when learning other-race faces.

Secondly, it is possible that whilst holistic and configural processing may be disrupted

6

for other-race faces compared to same-race faces, processing of facial motion may not be. Thus, observers may be able to utilize facial movement information to support recognition of other-race faces. Some studies have provided relevant, albeit indirect evidence to support this possibility. Lange et al. (2009) suggest that at least some aspects of the face and biological motion processing systems are dissociable with both common (e.g. configural processing) and separate (e.g., face specific and motion specific) processes involved. Additionally Bernstein and Yovel (2015) proposed a model of face processing that emphasizes the dissociation between form and motion as the primary functional division between the ventral stream - that goes through the fusiform face area (FFA) - and the dorsal stream - that goes through the superior temporal sulcus (STS). This position was also posited by Roark, Barrett, Spence, Abdi, and O'Toole (2003), who suggest that the neural networks involved in the structure-from-motion path converge with those involved in static face processing and that idiosyncratic facial movements are likely to be processed in the STS, separate from the normal face identity network (Haxby, Hoffman, & Gobbini, 2000; see also Pitcher, Dilks, Saxe, Triantafyllou, & Kanwisher, 2011). Similarly recent research (Pitcher, Duchaine & Walsh, 2014) has found that dynamic and static facial aspects are processed via dissociable cortical pathways. However, contrary to these findings Giese and Poggio's (2003) model of biological motion recognition posits that biological motion is processed in both the dorsal and ventral pathways, with information from both pathways then integrated in higher level areas like the STS. When applied to facial motion specifically, this model has received support from studies that have found dynamic faces to produce more activation in face sensitive areas such as the FFA and OFA (Fox, Iaria & Barton, 2009; Schultz, Brockhaus, Bülthoff & Pilz, 2013; Schultz & Pilz, 2009). But importantly it was also found that the FFA appears to process dynamic facial information differently from the STS (Schultz et al., 2013).

More direct evidence to support the idea that facial motion processing may not be

disrupted by face race was provided by Butcher, Lander, Fang and Costen (2011) who found that motion displayed by other-race faces at encoding lead to an improvement in recognition ability akin to that seen for same-race faces. Interestingly, Zhao, et al. (2014) found a significant other-race effect with static faces but not moving ones indicating that other-race face processing may benefit from the availability of motion information to such an extent that it eradicates the other-race effect.

However, the motion advantage for other-race faces is yet to be conceptually replicated to see whether it extends beyond standard recognition memory tasks. In the reported experiments a delayed visual search paradigm is used, replicating the basic methodology of Pilz et al. (2006), to explore the nature of the motion advantage for same- and other-race faces. Use of a delayed visual search task in the experiments reported here adds to our understanding of the motion advantage for same and other-race faces in a number of ways not associated with old/new recognition tasks. First, it can be considered an ecologically relevant perceptual task. The incidental learning phase ensures participants become familiar quickly with a limited number of target faces and the testing phase can be considered similar to "finding a friend in a crowd". Second, instead of making recognition judgements about a single face (i.e. have I seen this person before? yes / no), the visual search task requires participants to discriminate between different faces and identify which of the presented faces they recognise (i.e. I have seen that face before, that is [name]) which is arguably a more complex task. Third, manipulation of array size in the visual search task can offer insight to the mechanisms underlying the motion advantage. If the motion advantage is underpinned by differences in search efficiency we would expect to observe an interaction between presentation mode and array size. If no such interaction is observed it would indicate that the motion advantage reflects differences in decision criteria once the target face has been fixated. Finally, if a dynamic search advantage is found to generalise to other-race face learning it will establish the motion

advantage for same and other-race faces as a more robust effect that holds across experimental paradigms.

The aim of Experiment 1 specifically was to investigate whether facial motion leads to faster identification of familiarised same- and other-race faces (Pilz et al., 2006; 2009). It was proposed that facial motion at encoding would facilitate faster search latencies and that an other-race effect would be found. We hypothesised more efficient search performance for other-race faces (as well as same-race faces) learnt in motion than static.

Experiment 1

Method

Design. A within participants design was utilised with three independent variables manipulated; learning presentation mode (static and moving), race of face (British Caucasian i.e. same-race and Japanese i.e. other-race) and array size (2, 4 or 6 faces). Two dependent variables were measured assessing recognition accuracy; hits and reaction times (RTs) for correct responses.

Participants. 16 Caucasian participants (age range: 20-49; mean age: 26; 5 male, 11 female) with normal or corrected-to-normal vision participated in this experiment. All were students from the University of Manchester paid in participation credits and all were unfamiliar with the faces used as stimuli in the current experiment.

Stimuli and Apparatus. Face stimuli were selected from a bank of colour video sequences of British Caucasian and Japanese faces (previously used in Butcher et al., 2011; Lander, Hill, Kamachi & Vatikiotis-Bateson, 2007). All stimuli displayed at least the head and shoulders of

the subject with a blank, coloured backdrop and neutral expression. All were shot from a frontal position (1.1m from the camcorder) and were presented in full colour. Four male faces (two Caucasian, two Japanese) were selected to be used as target faces in the learning phase. Based on findings from a pilot study, the four faces were selected to be of average distinctiveness and were all matched on ratings of distinctiveness. For each target face, a static image and a 2s moving clip were used at learning. The moving clips displayed mainly non-rigid motion and all faces were seen speaking the same standardised sentence ("its fifteen minutes to three") although each spoke in their native language (note that clips were shown without audio). The static images were a single frame selected from the moving clip at a stage in the video when the face was displaying a neutral expression with no implied motion, by virtue of paused articulatory movements. In order to ensure that the image quality in the two presentation conditions was equivalent, the static images were not compressed when exported from the moving clip. It is therefore unlikely that any difference in recognition performance between the two learning presentation modes can be attributed to a difference in image quality. In both the static and moving conditions an ellipse shape and black background was placed around each target face at learning in order to reduce background cues, e.g. clothing, hair that could be used to identify the person rather than their face. Since the motion displayed in the moving clips was mainly non-rigid an equivalent amount of background information (e.g. hair) remained visible within the ellipse across moving and static conditions ensuring that the only difference between the two presentation modes was whether the face was seen in motion or not. The ellipse was applied at learning to ensure that during the later visual search task, participants were not able to use non face-based, low-level visual cues to recognition e.g. clothing colour.

For the recognition phase, a series of static visual search arrays (see figure 1) were constructed for use in the visual search task. Target present arrays contained one of the four

target faces plus one, three or five distracter faces. Target absent arrays contained two, four or six distracter faces. Within a single array each distracter face depicted a different person, with a total of 18 distracter faces utilised across the arrays (9 Caucasian, 9 Japanese). Faces displayed in each visual search array were of the same race and gender to control for possible race / gender categorisation effects that might confound search latencies. Previous studies have revealed an other-race categorisation advantage when participants are asked to categorise faces by their race (Levin, 1996; 2000; Valentine & Endo, 1992). For target present arrays, each target face was seen in all positions of the arrays at least twice. Overall this resulted in twenty-four 1 x 1 matrices (8 target absent and 16 target present arrays i.e. all four target faces shown in each of the two locations twice), forty-eight 2 X 2 matrices (16 target absent, 32 target present arrays i.e. all four target faces displayed in each of the four locations twice) and seventy-two 2 x 3 matrices (24 target absent, 48 target present i.e. all four target faces displayed in each of the 6 locations twice). Half of the arrays displayed same-race Caucasian faces and half other-race Japanese faces. The experiment was programmed using E-prime 2.0 software (Psychology Software Tools, Inc) and run on a Dell PC and 15" monitor.

Figure 1 about here

Procedure. The experiment consisted of two blocks (one for same-race and one for other-race faces) with block order counterbalanced. Within each block there was a learning and recognition phase. At learning participants were familiarised with the two target faces and were instructed that one was face A and the other face B. For each participant one target face in each block was presented as a static image and the other as a moving clip. Which of the two faces

11

was presented moving / static was counterbalanced across participants. Thus, any difference in search latencies for faces learnt in motion vs. static could not be attributed to differences in the recognisability of the faces, as target faces were used an equal number of times across presentation mode conditions. Target faces were displayed a total of 50 times each in the learning phase, alternating between the two, one at a time in the centre of the screen for 2s each with a 2s inter-stimulus interval (blank white screen). Participants were required to watch the faces carefully and complete a brief questionnaire about them (rating them on age, attractiveness, kindness and distinctiveness) to ensure that they engaged with the target faces. Following completion of the learning phase participants took a short break (approximately 3 minutes) before beginning the recognition task.

In the recognition phase participants completed a delayed visual search task. On each trial a fixation cross appeared in the centre of the screen followed by a 2, 4 or 6 static face array. Participants were instructed to decide as quickly and accurately as possible whether one of the target faces they had just learnt (face A or B) was present in the search array or not. If one of the targets was present they were to indicate which face it was, by pressing a corresponding key (i.e. they pressed one key if target face A was present, a second key if target face B was and a third key if no target was present). Each array remained onscreen until the participant made their response. In each block there was a total of 108 trials; thirty-six trials in each array size condition. In order to make the number of trials across array sizes equivalent, each 1 X 1 array was displayed three times and twelve of the 2 X 2 arrays were displayed twice. Across the two blocks participants completed a total of 216 visual search trials. Arrays were displayed in a random order within each block and target faces (A or B) were present on 66.7% of all trials. Each target face was present in the arrays an equal number of times. Recognition responses and RTs were recorded.

Results

The trials of interest were target present trials to which participants made correct recognition responses (hits). Reaction times (RTs) more than 2 standard deviations from the mean for that experimental condition were removed.

Hits analysis. The overall mean hit rate was 94.5%. Details of the hit rates exhibited across conditions are displayed in Figure 2. A 2 (learning presentation mode; moving, static) X 2 (race of face; same-, other-race) X 3 (array size; 2, 4 or 6) repeated measures ANOVA revealed no main effect of presentation mode, $F(1, 15) < .01$, $p = 1.00$, $\eta_p^2 > 0.001$, race, $F(1, 15) = .10$, $p = .75$, $\eta_p^2 = .007$, or array size, $F(2, 30) = 1.49$, $p = .24$, $\eta_p^2 = .09$. No interactions were found, all $F < 1.95$ and all $p > .16$.

Figure 2 about here

RT analysis. An overall mean reaction time (RT) of 1555ms was found. Figure 2 displays the RTs for each experimental condition. A 2 (learning presentation mode; moving, static) X 2 (race of face; same-, other-race) X 3 (array size; 2, 4 or 6) repeated measures ANOVA was performed. There was a main effect of learning presentation mode, $F(1, 15) = 4.54$, $p < .05$, $\eta_p^2 = .23$ revealing faster identification of faces learnt in motion ($M = 1508$ms, $SD = 266$ms) than faces presented in static at learning ($M = 1601$ms, $SD = 342$ms). A main effect of race was also found, $F(1, 15) = 24.81$, $p < .01$, $\eta_p^2 = .62$, with same-race targets identified faster ($M =$

1437ms, *SD* = 256ms) than other-race target faces (*M* = 1673ms, *SD* = 352ms). A main effect of array size was also found, $F(2, 30) = 94.52$, $p < .01$, $\eta_p^2 = .86$. Differences between the array size conditions were examined using post-hoc Bonferroni tests which found that search latencies for 2 face arrays (*M* = 1183ms, *SD* = 240ms) were significantly faster than for 4 face arrays (*M* = 1607ms, *SD* = 289ms; p < .01) and 6 face arrays (*M* = 1874ms, *SD* = 1530ms; p < .01) and that target faces were identified in 4 face arrays significantly faster than in 6 face arrays (p < .01). No interaction effects were found during analysis of RTs, all $F < 2.77$, all $p >$ .08 and all $\eta_p^2 < .16$. In particular there was no array size X learning presentation mode interaction i.e. no difference in the search slopes as a function of learning presentation mode, $F(2, 30) = 2.77$, $p = .08$, $\eta_p^2 = .16$.

Discussion

The results from Experiment 1 replicated the findings of Pilz et al. (2006, 2009) to establish a robust dynamic search advantage. RTs were faster for targets faces learnt in motion than static and increased as the number of faces in the array increased, but the effect of motion was not mediated by array size (as indicated by the absence of an array size X presentation mode interaction). This finding indicates that the dynamic search advantage is most likely the result of differences in decision criteria once the target has been located rather than differences in search efficiency. We return to this issue in the General Discussion.

Additionally, an other-race effect was observed in the overall RTs. Participants, here, as in previous studies (e.g. Bothwell et al., 1989), were faster to identify same-race target faces than other-race target faces. It is not possible to rule out differences due to the stimulus sets as an explanation of this other-race effect, as a cross race sample wasn't employed. However, what is of particular interest is the absence of an interaction between race and mode of presentation. This indicates that the strength of the dynamic search advantage did not differ

between same- and other-race faces. To our knowledge this finding provides the first evidence that the dynamic search advantage extends to other-race faces with observers benefiting from more robust representations of dynamically learnt other-race faces in a manner akin to same-race faces (Pilz et al., 2006; 2009). This finding together with Butcher et al. (2011) and Zhao et al. (2014) who used old/new recognition tasks, suggests that motion information exhibited by other-race faces at encoding can facilitate more accurate and faster recognition of same- and other-race faces.

In Experiment 1, accuracy was not influenced by either race or learning presentation mode as expected (e.g. race, Meissner & Brigham, 2001; presentation mode, Lander & Bruce, 2003). Differences in accuracy for static compared to moving faces did not reach significance in Pilz et al. (2006) and only reached significance in Pilz et al. (2009) in one block of Experiment 1. It is possible that in all these studies accuracy results may have been limited by ceiling effects since here the condition with the lowest accuracy (same-race static, 6 face arrays) was still highly accurate ($M = 91\%$). These findings are not surprising given observers were asked to learn just four faces overall. Learning just two faces in each block ensured that observers were quickly familiarised with the target faces, shifting importance away from accuracy, onto reaction time. Therefore, the question remains whether the dynamic search advantage for both same- and other-race faces can be observed in accuracy as well as RTs.

In Experiment 2 we increase task difficulty to investigate whether target same- and other-race faces familiarised with facial movement facilitate more accurate, as well as faster, face identification. Firstly, we increased the number of familiarised faces across the experiment, by including two additional blocks for female same- and other-race faces. Secondly, the task was made more difficult by cropping all recognition phase images to remove all external cues and just display the face.

Additionally, in order to further explore explanations of how facial movement influences performance for both same- and other-race faces, Experiment 2 investigated whether the effect of motion on visual search latencies is dependent on the number of times the participant views each target face and / or the amount of motion information they observe for each target face. In Experiment 1, target faces were viewed 50 times and each time were presented for 2s whereas Pilz et al. (2006) presented each target face 100 times, each time for 1040ms. Although overall exposure time to each target face was approximately equivalent (~100s) across the studies and both demonstrated a dynamic search advantage, these learning conditions differ in two potentially important ways; a) face repetitions and b) amount of motion information observed (as a result of the clip length). In Experiment 2 we investigated the dynamic search advantage for same- and other-race faces across three learning conditions. The first condition (Condition 1) was analogous to the learning procedure in Experiment 1 with 50 repetitions of a 2s clip of each face. The second condition (Condition 2) was analogous to that used by Pilz et al. (2006) with 100 repetitions of a 1s clip per face. In the third condition, as in Condition 1, participants saw 50, 2s repetitions of each face but rather than see the full 2s clip used in Condition 1, here the 1s clip used in Condition 2 was repeated twice on each repetition. So, Condition 3 differed to Condition 2 in the number of face repetitions but displayed the same amount of motion information as Condition 2. Likewise Condition 3 differed to Condition 1 in the amount of motion information displayed but maintained the same number of repetitions as Condition 1. If the number of facial motion viewing episodes is fundamental to the dynamic search advantage then we would expect differences in the motion advantage gained between Conditions 2 and 3. On the other hand if the amount of facial motion information observed is essential to the dynamic search advantage we should expect to find a difference in the motion advantage gained between Conditions 1 and 3.

Experiment 2

Method

Participants. 90 Caucasian participants (37 male, 53 female) between 19 and 25 years old (mean 21 years and 3 months) participated in this experiment. All were students from the University of Manchester paid in participation credits. All had normal or corrected to normal eyesight and had not taken part in any experiments of this kind previously.

Design. A mixed design was utilised with four independent variables; learning presentation mode (static and moving; within), learning condition (condition 1, condition 2 or condition 3; between), race of face (British Caucasian and Japanese; within) and array size (2, 4 or 6 faces; within). We did not include face gender as a factor as our initial analysis demonstrated no significant effect of gender, thus, we collapsed across gender. Hits and reaction times (RTs) for correct responses were measured.

Stimuli and Apparatus. The face stimuli used in Experiment 1 were again utilised except that an additional 4 targets faces (two female Caucasian, two female Japanese) were now selected to act as additional target stimuli resulting in a total of 8 target faces in Experiment 2. For each target face a static image, a 2s movie clip and a 1s movie clip were used as learning stimuli across three different learning conditions, as follows;

Learning Condition 1 - each face was shown 50 times for two seconds. In this condition the moving face was a full two second clip of the face speaking the full sentence with an inter-stimulus interval of 50ms during which a white screen with a black fixation cross in the centre was displayed.

Learning Condition 2 – each face was shown 100 times for one second with an inter-stimulus interval of 50ms during which a white screen with a black fixation cross in the centre was displayed.

Learning Condition 3 – each face was shown 50 times for two seconds. In this condition the two second moving clip was the one second clip from condition two repeated twice, giving the same amount of visual movement as condition two but on less occasions. Again, the presentations were separated by the 50ms fixation cross.

The visual search arrays used in Experiment 1 were again utilised. An additional 144 arrays were constructed using the same method as in Experiment 1 but containing female same- and other-race faces; twenty-four 1 x 1 matrices (8 target absent and 16 target present arrays i.e. all four target faces shown in each of the two locations twice), forty-eight 2 X 2 matrices (16 target absent, 32 target present arrays i.e. all four target faces displayed in each of the four locations twice) and seventy-two 2 x 3 matrices (24 target absent, 48 target present i.e. all four target faces displayed in each of the 6 locations twice). Half of these additional arrays displayed same-race Caucasian female faces and half other-race Japanese female faces. In Experiment 2 each recognition phase image had the face area cropped out and placed against a plain black background to eliminate any external features such as clothes or background, which may act as a cue to recognition.

Procedure. The experiment consisted of two phases, learning and recognition that were repeated for each block (Caucasian male, Caucasian female, Japanese male, Japanese female; counterbalanced ordering across participants). The learning and recognition phase procedure was the same as in Experiment 1 except that observers were familiarised with the target faces in the three conditions outlined above. For each of the four blocks there were 108 trials (thirty-

six 1 X 1, thirty-six 2 X 2 and thirty-six 2 X 3). In total participants completed 432 trials. Arrays were displayed in a random order for each participant and target faces (A or B) were present on 66.7% of all trials with the remaining 33.3% target absent trials. Each target face was present in the arrays an equal number of times.

Results

Analysis was carried out on target present trials to which participants gave correct responses.

Hits analysis. The overall mean hit rate was 79.8%. See Figure 3 for full details of the hits rates in each experimental condition.

Figure 3 about here

A 2 (learning presentation mode; moving & static) X 3 (learning condition; condition 1, condition 2 or condition 3) X 2 (race of face; same- and other-race) X 3 (array size; 2, 4 or 6 faces) ANOVA was conducted. There was a significant main effect of learning presentation mode, $F(1,87) = 4.35$, $p < .05$, $\eta_p^2 = .05$, with more hits when learning moving ($M = 81.2\%$, $SD = 31.1\%$) compared with static ($M = 78.5\%$, $SD = 18.0\%$) faces. There was no significant main effect of learning condition, $F(2,87) = 2.93$, $p > .05$, $\eta_p^2 = .06$. A significant main effect of race was found, $F(1,87) = 20.82$, $p < .05$, $\eta_p^2 = .19$. Performance was better for same-race compared ($M = 82.9\%$, $SD = 17.4\%$) with other-race faces ($M = 76.8\%$, $SD = 19.7\%$). There was a significant main effect of array size, $F(2,174) = 7.29$, $p < .01$, $\eta_p^2 = .08$. Differences

between the three array size conditions were examined using post-hoc Bonferroni tests which found that search latencies for 2 face arrays ($M = 82.2\%$, $SD = 19.4\%$) were significantly faster than for 4 face arrays ($M = 80.2\%$, $SD = 19.2\%$; p < .05) and 6 face arrays ($M = 76.8\%$, $SD = 17.4\%$; p < .01) and that target faces were identified in 4 face arrays significantly faster than in 6 face arrays (p < .01). Finally, no interactions were found, all $F < 1.90$ and all $p > .18$.

RT analysis. The overall mean RT was 1562ms. Figure 3 shows the mean reaction times and for each experimental condition. A 2 (learning presentation mode; moving & static) X 3 (learning condition; condition 1, condition 2 or condition 3) X 2 (race of face; same- and other-race) X 3 (array size; 2, 4 or 6 faces) ANOVA was conducted. This revealed a significant main effect of learning presentation mode, $F(1,87) = 12.05$, $p < .01$, $\eta_p^2 = .12$ with faster response times when learning moving ($M = 1505$ms, $SD = 648$ms) compared with static ($M = 1619$ms, $SD = 647$ms) faces. There was no significant main effect of condition, $F(2,87) = 1.52$, $p > .05$, $\eta_p^2 = .03$. However there was a significant main effect of race, $F(1,87) = 40.52$, $p < .01$, $\eta_p^2 = .32$. Identification was significantly faster for same-race compared ($M = 1454$ms, $SD = 693$ms) with other-race faces ($M = 1670$ms, $SD = 533$ms).   There was also a significant main effect of array size, $F(2,174) = 22.67$, $p < .01$, $\eta_p^2 = .21$. Differences between the three array size conditions were examined using post-hoc Bonferroni tests, which found that search latencies for 2 face arrays ($M = 1407$ms, $SD = 539$ms) were significantly faster than for 4 face arrays ($M = 1540$ms, $SD = 728$ms; p < .05) and 6 face arrays ($M = 1739$ms, $SD = 564$ms; p < .01) and that target faces were identified in 4 face arrays significantly faster than in 6 face arrays (p < .01). Finally, there was a significant two way interaction between race and array size, $F(2,174) = 3.33$, $p <.05$, $\eta_p^2 = .04$. Analysis of the simple main effect showed that there was a larger RT difference between same- and other-race faces when the array size was two (*p*

< .001) or four ($p <.001$) compared with 6 ($p < .05$). No other interactions were found, all $F <$ 1.50 and all $p > .20$ and all $\eta_p^2 < .045$ for non-significant results.

Discussion

As in Experiment 1 we found faster search latencies for a) target faces familiarised in motion compared to static b) same-race targets compared to other-race targets and c) smaller arrays compared to larger arrays. These reaction time effects were mirrored in the accuracy analysis suggesting that task difficulty may account for the absence of a motion advantage and other-race effect in the accuracy data for Experiment 1. Importantly, with increased task difficulty we found a dynamic search advantage for both same- and other-race faces in accuracy as well as reaction times. Previous studies have failed to observe the effect in accuracy (Pilz et al., 2006) or have not found it consistently across experimental blocks (Pilz et al., 2009). In order to reconcile these inconsistent accuracy findings a flexible usage criterion is hypothesised in which the visual system mediates the use of identity specific information stored in any given face representation based upon the necessity of that information to accomplishing the task at hand (Morrison & Schyns, 2001; Olivia & Schyns, 1997; Schyns & Olivia, 1999). It is argued that, as has been demonstrated in scene and object perception, a selective use of diagnostic information may be evident in face processing. When recognising faces observers may not use all the information available to them, but instead select the most diagnostic elements for the task at hand. The visual system is capable of selectively extracting this information from the visual array in order to perform flexible recognition decisions dependent on the task (Schyns & Gosselin, 2003). As such, with increased task difficulty in Experiment 2 the reliance on motion as a cue to identity may have increased compared to Experiment 1 and other studies that have found no difference between static and moving conditions in terms of accuracy (e.g. Pilz et al., 2006).

Crucially there was no main effect of learning condition here, nor did learning condition interact with presentation mode on either the RT or accuracy analysis. This finding suggests that the motion advantage in face recognition, observed here using a delayed visual search task, is relatively robust and consistent regardless of the number of repeat exposures to a target face and the amount of motion information displayed (as a function of clip length).

General Discussion

The experiments reported here allow us to make the following points.

First, we have demonstrated a clear dynamic search advantage; learning a face in motion leads to significantly faster and more accurate visual search performance than learning from a static face. Thus, we have replicated the RT findings of Pilz et al. (2006, 2009) who also used a delayed visual search task to explore the motion advantage. Importantly, we also observed a dynamic search advantage in recognition accuracy (Experiment 2). This finding indicates that the absence of a motion advantage in the accuracy analysis of both Experiment 1 and Pilz et al. (2006) may be explained by differential task difficulty. In Experiment 2 we increased task difficulty by cropping external cues at recognition and requiring participants to learn more faces. In doing so we support the notion that the delayed visual search paradigm offers a useful task when investigating the motion advantage in face recognition. However, further systematic investigation is needed to explore why some tasks reveal a motion advantage, while others do not.

Interestingly, use of the visual search paradigm offers further insight into the nature of the motion advantage. Although significant effects of array size were found, with accuracy and RTs decreasing / increasing respectively as a function of array size, faces learnt in motion and static were equally effected by array size. If the dynamic search advantage is underpinned by

differences in search efficiency we would expect to observe an interaction between presentation mode and array size. This interaction was not observed here or in Pilz et al. (2006), indicating that the dynamic search advantage reflects differences in decision criteria once the target face has been fixated. This finding can be explained by the representation enhancement hypothesis (O'Toole et al., 2002), which suggests that seeing a face move at encoding leads to the construction of more robust face representations, as motion aids the perception of three-dimensional structure. These robust representations then facilitate identification of the face across different viewing conditions and contexts. Indeed it is possible that face representations encoded in motion are themselves dynamic in nature (i.e. a '*dynamic mental representation*' Freyd, 1987) capturing both the form and motion characteristics of the face. This notion has been supported by Thornton and Koutzi (2001), Pilz et al., (2006; 2009) and Pilz, Vuong, Bülthoff and Thornton (2011) and is also ecologically plausible given that the visual system is constantly exposed to faces in motion rather than static. Thus, it is likely that the neural mechanisms for encoding dynamic stimuli might be more robust than those for encoding evolutionary-rare static face stimuli.

Additionally, it is unlikely that the beneficial effect of movement seen here was simply due to the increased amount of static information within the dynamic stimuli. Results have, to date, provided no evidence to support this explanation of the motion advantage as presentation of multiple static images of a face is not seen to produce the same level of recognition performance as moving conditions (Christie & Bruce, 1998; Lander et al., 1999; Pike et al., 1997). These findings suggest that actually viewing the face in motion is fundamental to this effect.

Second, we have demonstrated that the motion advantage, and more specifically the dynamic search advantage, extends to other-race faces. Previous literature on the motion advantage in face recognition has almost exclusively used same-race stimuli despite the

importance of race in face recognition (Bothwell Brigham & Malpass, 1989; Meissner & Brigham, 2001). The results from this study offer the first demonstration that the dynamic search advantage can be generalized to other-race faces and support recent research (Butcher et al., 2011) that used a standard yes/no recognition task and found that seeing a face move non-rigidly aids the learning of other-race faces, as well as same-race faces. This finding has since been supported by Zhao et al. (2014). As a result, it is now clear that across both yes/ no recognition *and* visual search tasks, observers are able to use facial motion information exhibited by other-race faces at encoding as an aid to recognition, as they have been seen to do previously for same-race faces (Knappmeyer et al., 2003; Lander & Bruce, 2003; Thornton & Kourtzi, 2002). These results suggest that like same-race faces, other-race faces also benefit from the development of more robust mental representations as a result of being viewed in motion at encoding (O'Toole et al., 2002). They also suggest that despite perceptual processing abilities (e.g. configural and holistic processing) being argued to be weaker for other-race faces (e.g. Tanaka et al., 2004), other-race observers do not lack the perceptual expertise to process and utilize other-race facial motion to aid recognition. In addition to the general explanations of the motion advantage (discussed above), motion may be enhancing recognition of other-race faces, in particular, in two ways. First, it is possible that the availability of facial motion at learning facilitates the holistic and / configural processing that we know to be a hallmark of same-race face processing (e.g. Young, Hellawell, & Hay, 1987). It has been claimed that same-race faces are processed more holistically than other-race faces (Rossion & Michel, 2011, for a review) and that there is an own-race advantage in processing configural information (e.g. Rhodes, Hayward & Winkler, 2006). Thus, a shift towards greater holistic or configural processing, as a result of viewing a face in motion (Xiao et al., 2015), may enable observers to process other-race faces in a similar way to same-race faces, which typically produce better recognition performance. Alternatively, Pitcher et al. (2014) found there is a dissociation in the

cortical pathways for dynamic and static faces and it has also been found that at least some aspects of the face and biological motion processing systems are dissociable with both common (e.g. configural processing) and separate (e.g., face specific and motion specific) processes involved (Lange et al., 2009). These earlier findings offer the possibility that the enhanced performance seen here for other-race faces learnt in motion compared to static might, in part, be explained by motion specific systems not being effected by race to the extent that common processes, like configural processing, are argued to be. A recent demonstration that the other-race effect is sensitive to presentation mode (Zhao et al., 2014) also supports this explanation of the beneficial effect of motion as it indicates that motion specific systems may not be effected by race to the extent that common processes are, enabling motion to serve as a cue to identity recognition for other-race faces. However, others have found evidence opposing the cortical dissociation of form and motion that is fundamental to this notion (see Fox et al., 2009; Schultz & Pilz, 2009; Schultz et al., 2013). Additionally, it is worth noting that in the current findings race and presentation mode did not interact indicating that the other race effect was observed in our moving condition too. Further work is needed to explore the importance of holistic, configural and featural processing and motion at encoding on the other-race effect.

Third, we have demonstrated that regardless of the number of repeat exposures to a face and the amount of motion information displayed (as a function of clip length) during familiarisation, the dynamic search advantage for same- and other-race face identification is relatively robust, with different learning conditions exhibiting a similar effect size. This is interesting in light of previous inconsistent findings in the motion advantage literature when other experimental paradigms have been employed (Christie & Bruce, 1998; Lander & Bruce, 2003; Pike et al., 1997; Thornton & Kourtzi, 2002). Furthermore, these findings offer greater understanding of the parameters under which the dynamic search advantage occurs for both same- and other-race faces, revealing that the effect does not require a specific number of face

repetitions or amount of motion information. However, it remains to be seen whether the size of this dynamic search advantage is related to total face exposure time since here the *total* duration of movement was the same across all three learning conditions (Experiment 2). It is possible that reducing the total presentation time for moving targets might reduce the dynamic advantage. However, Lander and Davies (2007) revealed that observers are able to extract and encode dynamic information even when viewing very short moving clips. Here, again a dynamic search advantage was observed despite observers being exposed to very short moving sequences. These findings suggest that the motion advantage may not be dependent on exposure time. Instead it may be that observers rapidly learn characteristic motion patterns exhibited by a face to aid later recognition. However, given that we have also previously found that the magnitude of the motion advantage for famous faces is significantly correlated with face familiarity (Butcher & Lander, 2016) future research is required to investigate the strength of the dynamic search advantage when familiarity, as a result of exposure time, is systematically varied. This would help to determine whether the effect is related to the availability of more information or indeed that the motion advantage reflects the fact that facial motion facilitates the encoding of more robust facial representations at encoding for both same and other-race faces.

In conclusion, the current findings provide evidence that the dynamic search advantage extends to other-race faces. This offers converging evidence that motion depicted by other-race faces can lead to faster reaction times and greater identification accuracy, analogous to same-race faces (Butcher et al., 2011; Zhao et al., 2014). However, unlike Zhao et al. (2014) our findings indicate that whilst recognition performance for other-race faces may benefit from the availability of motion information, the other race effect is not restricted to a static mode of presentation. Finally, using an incidental learning and visual search methodology we have demonstrated that the effect of motion on the recognition of both same- and other-race faces is

relatively robust and not restricted to situations where encoding involves seeing the face a particular number of times or displaying a particular amount of motion information.

References

Bernstein, M., & Yovel, G. (2015). Two neural pathways of face processing: A critical

evaluation of current models. *Neuroscience and Biobehavioural Reviews, 55*, 536-46.

Bothwell, R. K. Brigham, J. C., & Malpass, R. S. (1989). Cross-racial identification.

Personality and Social Psychology Bulletin, 15, 19-25.

Braje, W.L., Kersten, D., Tarr, M. J., & Troje, N. F. (1998) Illumination effects in face

recognition. *Psychobiology, 26,* 271-380.

Butcher, N. & Lander, K. (2016). Exploring the motion advantage: evaluating the

contribution of familiarity and differences in facial motion. *Quarterly Journal of

Experimental Psychology,* DOI: 10.1080/17470218.2016.1138974

Butcher, N., Lander, K., Fang, H., & Costen, N. (2011). The effect of motion at encoding and

retrieval for same and other race face recognition. *British Journal of Psychology, 102,* 931-

942.

Christie, F., & Bruce, V. (1998). The role of dynamic information in the recognition of

unfamiliar faces. *Memory & Cognition, 26,* 780-790.

Fox, C. J., Iaria, G., & Barton, J. J. S. (2009). Defining the face processing network;

optimization of the functional; localizer in fMRI. *Human Brain Mapping, 30*(5), 1637-

1651.

Freyd, J. J. (1987). Dynamic mental representations. *Psychological Review, 94*(4), 427-438.

Giese, M. A., & Poggio, T. (2003). Neural mechanisms for the recognition of biological

movements. *Nature Reviews Neuroscience,* 4(3), 179-192.

Harrison, S. A., Richler, J. J., Mack, M. L., Palmeri, T. J., Hayward, W. G., & Gauthier, I.

(2011). The complete design lets you see the whole picture: Differences in holistic

processing contribute to face-inversion and other-race effects. *Journal of Vision, 11*, 625.

Hayward, W. G., Crookes, K., & Rhoves, G. (2013). The other-race effect: Holistic coding differences and beyond. *Visual Cognition, 21*, 1224-1247.

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

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

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

Hill H, Schyns PG, Akamatsu S. (1997). Information and viewpoint dependence in face recognition. *Cognition, 62,* 201-22.

Hugenberg, K., Young, S. G., Bernstein, M. J., & Sacco, D. F. (2010). The categorization-individuation model: an integrative account of the other race recognition deficit. *Psychological Review, 117,* 1168-1187.

Johnson, K. J., & Fredrickson, B. L. (2005). "We all look the same to me": Positive emotions eliminate the own-race bias in face recognition. *Psychological Science, 16,* 875–881.

Knappmeyer, B., Thornton, I.M., & Bülthoff, H. H. (2003). Facial motion can bias the perception of facial identity. *Vision Research, 43,* 1921-1936.

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

Lander, K., & Bruce, V. (2003). The role of motion in learning new faces. *Visual Cognition, 10,* 897-912.

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

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.

Lander, K., Hill, H., Kamachi, M., & Vatikiotis-Batson, E. (2007). It's not what you say but the way you say it: Matching faces and voices. *Journal of Experimental Psychology: Human Perception and Performance, 33,* 903-914.

Lange, J., de Lussanet, M., Kuhlmann, S., Zimmermann, A., Lappe, M., Zwitserlood, P., & Dobel., C. (2009). Impairments of biological motion perception in congenital prosopagnosia. *PLos ONE, 4(10)*, e7414. doi:10.1371/journal.pone.0007414

Levin, D. T. (1996). Classifying faces by race: The structure of face categories. *Journal of Experimental Psychology-Learning Memory and Cognition, 22*, 1364-1382.

Levin, D. T. (2000). Race as a visual feature: Using visual search and perceptual discrimination tasks to understand face categories and the cross-race recognition. *Journal of Experimental Psychology-General, 129*, 559-574.

McKone, E. (2004). Isolating the special component of face recognition: Peripheral identification and a Mooney face. *Journal of Experimental Psychology: Learning, Memory, and Cognition 30,* 181 – 197.

Meissner, C. A., & Brigham, J. (2001). Thirty years of investigating the own-race bias in face recognition: A meta-analytic review. Paper presented at the meetings of the American Psychology-Law Society, New Orleans, LA.

Meissner, C. A., Brigham, J. C., & Butz, D.A. (2005). Memory for own- and other-race faces: A dual-process approach. *Applied Cognitive Psychology, 19,* 545-567.

Morrison, D. J., & Schyns, P. G. (2001).Usage of spatial scales for the categorization of faces, objects, and scenes. *Psychonomic Bulletin & Review. 8*, 454 - 469.

Oliva, A., & Schyns, P. G. (1997). Coarse blobs or fine edges? Evidence that information diagnosticity changes the perception of complex visual stimuli. Cognitive Psychology, 34, 72–107.

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

Pike, G. E., Kemp, R. I., Towell, N., & Phillips, K. C. (1997). Recognizing Moving Faces: The Relative Contribution of Motion and Perspective View Information. *Visual Cognition, 4,* 409-438.

Pilz, S. K., Bülthoff, H. H., & Vuong, Q. C. (2009). Learning influences the encoding of static and dynamic faces and their recognition across different spatial frequencies. *Visual Cognition, 17,* 716-735.

Pilz, K. S., Thornton, I. M., & Bülthoff, H. H. (2006). A search advantage for faces learned in motion. *Experimental Brain Research, 171,* 436-447.

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.

Pitcher, D., Dilks, D. D., Saxe, R. R., Triantafyllou, C., & Kanwisher, N. (2011). Differential selectivity for dynamic versus static information in face selective cortical regions. *NeuroImage, 56,* 2356-2363.

Pitcher, D., Duchaine, B., & Walsh, V. (2014). Combined TMS and fMRI reveals dissociable pathways for dynamic and static face perception. *Current Biology, 24,* 2066-2070.

Rhodes, G., Hayward, W. G., & Winkler, C. (2006). Expert face coding: Configural and component coding of own-race and other-race faces. *Psychonomic Bulletin & Review, 13*, 499–505.

Rhodes, G., Tan, S., Brake, S., & Taylor, K. (1989). Expertise and configural coding in face recognition. *British Journal of Psychology, 80,* 313-331.

Roark, D. A., Barrett, S. E., Spence, M. J., Abdi, H., & O'Toole, A. J. (2003). Psychological

and neural perspectives on the role of motion in face recognition. *Behavioral and Cognitive Neuroscience Reviews, 2,* 15-46.

Rossion B., & Michel C. (2011). An experience-based holistic account of the other-race face effect. In G. Rhodes., A. Calder., M. Johnson., J. V. Haxby (Eds.), *Oxford Handbook of Face Perception* (pp.215-244). Oxford: Oxford University Press.

Schultz, J., Brockhaus, M., Bülthoff, H. H., & Pilz, K. (2013). What the Human Brain Likes About Facial Motion. *Cerebral Cortex, 23*(5), 1167-1178.

Schultz, J., & Pilz, K. S. (2009). Natural facial motion enhances cortical responses to faces *Experimental Brain Research, 194*(3), 465-475.

Schyns, P. G., & Gosselin, F. (2003). Diagnostic use of scale information for componential and holistic recognition. In M, A. Peterson, & G. Rhodes (Eds.), *Perception of Faces, Objects, and Scenes: Analytic and Holistic Processes Advances in Visual Cognition* (pp. 120 – 148). New York: Oxford University Press.

Schyns, P. G., & Oliva, A. (1999). Dr. Angry and Mr. Smile: When categorization flexibly modify the perception of faces in rapid visual presentation. *Cognition, 69***,** 243–265.

Tanaka, J. W., Kiefer, M., & Bukach, C. M. (2004). A holistic account of the own-race effect in face recognition: Evidence from a cross-cultural study. *Cognition, 93,* B1–B9.

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

Thornton, I. M., Mullins, E., & Banahan, K. (2011). Motion can amplify the face-inversion effect. *Psihologija, 44*, 5-22.

Troje, N. F., Bülthoff, H. H. (1996). Face Recognition under varying poises: the role of texture and shape. *Vision Research, 12,* 1761-1771.

Valentine, T. & Bruce, V. (1986). Recognizing familiar faces: The role of distinctiveness and familiarity. *Canadian Journal of Psychology, 40,* 300-305.

Valentine, T., & Endo, A. (1992). Towards an exemplar model of face processing; The effect of race and distinctiveness. *Quarterly Journal of Experimental Psychology; Human Experimental Psychology, 44a,* 671-703.

Wagenaar, W. A., & van der Schrier, J. H. (1996). Face recognition as a function of distance and illumination: A practical tool for use in the courtroom. *Psychology, Crime & Law, 2,* − 332.

Wheeler, A., Anzures, G., Quinn, P. C., Pascalis, O., Omrin, D. S., & Lee, K. (2011). Caucasian infants scan own- and other-race faces differently. *PLoS ONE, 6*, e18621.

Xiao, N. G., Quinn, P.C., Ge, L., & Lee, K. (2012). Rigid facial motion influences featural, but not holistic, face processing. *Vision Research, 57*, 26-34.

Xiao, N. G., Quinn, P.C., Ge, L., & Lee, K. (2013). Elastic facial movement influences part-based but not holistic processing. *Journal of Experimental Psychology: Human Perception and Performance, 39*, 1457-1467.

Xiao, N. G., Quinn, P.C., Liu, S., Ge, L., Pascalis, O., & Lee, K. (2015). Eye tracking reveals a crucial role for facial motion in recognition of faces by infants. *Developmental Psychology, 51*, 744-757.

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

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configural information in face perception. *Perception, 16,* 747-759.

Zhao, M., Hayward, W., & Bulthoff, I. (2014). Face format at encoding affects the other-race effect in face memory. *Journal of Vision, 14*(6), 1-13.

Figure 1. Example of a 2 X 2 other-race male array used in the visual search task (Experiment 1).

Figure 2. Accuracy (top) and Reaction time (bottom) data from Experiment 1 for each target race (same-race, other-race) and presentation mode (static, moving).

Figure 3. Accuracy (top) and Reaction time (bottom) data from Experiment 2 for each target race (same-race, other-race) and presentation mode (static, moving).