Learning New Faces From Multiple Images

Chris Longmore[1], Isabel M. Santos[2], Mariana Carrito[2], Ana Pereira[2], Paulo Rodrigues[3] and Carlos F. Silva[2]

[1]University of Plymouth, UK, [2]University of Aveiro, Portugal, [3]University of Beira Interior, Portugal

Email: chris.longmore@plymouth.ac.uk Twitter: @ChrisALongmore

Introduction

It has been demonstrated that the recognition of unfamiliar faces (e.g. faces that have been seen only once) is poor. Furthermore, recognising an unfamiliar face from an image different to one already seen (e.g. by a change in viewpoint) is particularly challenging (Bruce, 1982) even when the studied image has been learnt extensively (Longmore, Liu & Young, 2008). Longmore et al. suggest that in such studies, participants learn properties of the **image** as opposed to invariant structural properties of the **face** - limiting generalisation to novel images of a face. In their study, Longmore et al. required participants to interpolate between the full-face and profile views of a face in order to achieve successful recognition at a previously unseen three-quarter view - a task that was demonstrated to be sufficiently difficult that performance was no better after learning two views of a face than after learning one view (see Figure 1).

It is possible that viewpoint changes are sufficiently extreme that obtaining an advantage for learning two views over a single view is difficult as overlap between the two images is small. Two experiments were conducted to investigate whether

a) maintaining the viewpoint of the learnt and test views and using another image transformation (age) aids generalisation to novel images of a newly learnt face
b) neurocorrelates of generalisation could be established using ERPs.

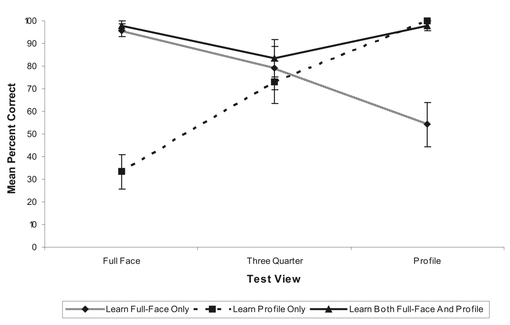

Figure 1: The results of Longmore et al. (2008), experiment 3. Participants performance on a three-quarter view at test does not benefit from learning two view of the face over a single view.

Experiment 1 - Generalisation Across Ages

Participants

Fifteen participants took part in Experiment 1.

Procedure

Participants were trained to recognise 12 individuals by name; 4 from a 20 year old face only, 4 from a 60 year old face only and 4 from both the 20 and 60 year old images (see Figure 2). Training was complete when participants could identify the individuals by name without making an error.

Figure 2: Images used in Experiment 1. From left to right; 20 year old, 40 year old (morphed from the 20 year old and 60 year old images) and 60 year old.

Recognition accuracy was assessed for three images of each person; both the original 20 year old and 60 year old images together with a 50% morph between these two ages to give a pseudo 40 year image.

Results

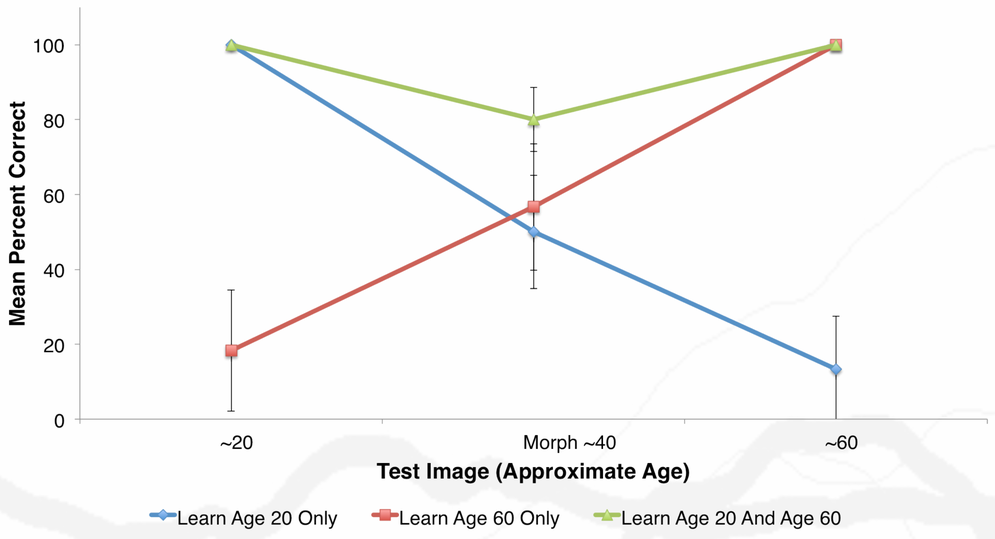

Figure 3: Average performance on each of the three test views for the three learning conditions. Error bars represent 95% confidence intervals.

Recognition accuracy for the morphed 40 year old image after learning both 20 year old and 60 year old images was significantly poorer than for either of the learnt views; $F(2,28) = 8.034$, $p = .02$. Learning two views did however aid recognition of the morphed 40 year old image in comparison to learning only a single image; $F(1,14) = 21.00$, $p < .01$. The results suggest that whilst recognition is still heavily reliant on studied images, generalisation can be aided by studying two views of a face.

Experiment 2 - Evidence from ERPs

Objective

Investigate the electrophysiological correlates of learning new faces using a paradigm similar to Longmore et al (2008). Test whether it is possible to establish the neurocorrelates of any generalisation that takes place when learning new faces on two different viewpoints (frontal and profile), which would facilitate recognition on an unseen three-quarter view, as compared to learning the faces on a single viewpoint.

Participants

Nineteen right-handed university students took part.

Materials

Training set: 16 facial images of 12 different individuals, all male (4 in full-face view, 4 in profile view, and 4 both in full-face and profile views). Test set: Same 12 individuals, each displayed in full-face, profile and three-quarter views (target faces) and 12 additional individuals, each displayed in the three mentioned viewpoints (distractor faces). See Figure 4 for an example of the images used in the test phase.

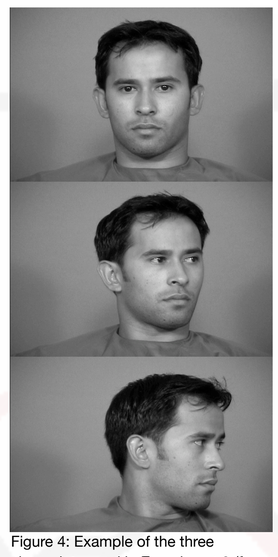

Figure 4: Example of the three viewpoints used in Experiment 2 (from top to bottom): full-face, three-quarter and profile.

Procedure

Same procedure as Experiment 1, except that the learning phase ended after 10 complete rounds of presentation of the 16 facial images and respective identification attempts by the participants (number decided based on previous results). In the test phase, the 72 stimuli were randomly presented and repeated 6 times in a blocked manner, in order to increase the signal to noise ratio for the ERP analyses.

Results

Figure 5: Average waveforms on the examined electrodes relative to the effect of number of learning views.

There was a marginally significant main effect of number of learning views on the wave N170, $F(1,18) = 4.22$, $p = .055$, with faces learnt in a single view eliciting more negative amplitudes ($M = -7.12$ μV) than the faces learnt in two viewpoints ($M = -6.12$ μV) when the test view had been previously learnt. Also for the N170, in the analysis of the effect of number of learning views on the recognition in the unseen three-quarter viewpoint, there was a marginally significant interaction between the number of learning views and location, $F(2,36) = 3.14$, $p = .055$. Pairwise comparisons revealed that, for the occipital electrodes, faces learnt in a single view elicited significantly more negative amplitudes ($M = -8.59$ μV) than faces learnt in two viewpoints ($M = -7.42$ μV).

Conclusions

The results of Experiment 1 suggest that whilst generalisation to novel views of a face is possible, performance is still below that for learnt images. The ERP data from Experiment 2 are however compatible with a facilitation effect of learning two viewpoints when recognition on a previously unseen view is required compared to learning a single viewpoint, indicating some degree of generalisation from full-face and profile viewpoints to an unseen three-quarter view. This suggests that some kind of structural information is being extracted during learning that facilitates recognition from novel views, even though this is not yet apparent in behavioural performance.

References

Bruce, V. (1982). Changing faces - Visual and non-visual coding processes in face recognition. *British Journal of Psychology*, *73*, 105-116

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

Research funded by