Article:

Ritchie, Kay Laird and Burton, A. Mike orcid.org/0000-0002-2035-2084 (2017) Learning faces from variability. The Quarterly Journal of Experimental Psychology. pp. 897-905. ISSN: 1747-0226

https://doi.org/10.1080/17470218.2015.1136656

Learning faces from variability

Journal:	*Quarterly Journal of Experimental Psychology*
Manuscript ID	QJE-SIP 15-210.R1
Manuscript Type:	Special Issue Paper
Date Submitted by the Author:	n/a
Complete List of Authors:	Ritchie, Kay; University of York, Burton, Mike; University of York, ;
Keywords:	face learning, variability, face recognition

Learning faces from variability

Kay L. Ritchie & A. Mike Burton

Department of Psychology, University of York, UK

Corresponding author: Mike Burton, Department of Psychology, University of York, York, YO10 5DD, England, United Kingdom. Email: mike.burton@york.ac.uk

Key words: Face learning, face recognition, variability.

Abstract

Research on face learning has tended to use sets of images which vary systematically on dimensions such as pose and illumination. In contrast, we have proposed that exposure to naturally varying images of a person may be a critical part of the familiarisation process. Here, we present two experiments investigating face learning with "ambient images" – relatively unconstrained photos taken from internet searches. Participants learned name and face associations for unfamiliar identities presented in high or low within-person variability, i.e. images of the same person returned by internet search on their name (high variably) versus different images of the same person taken from the same event (low variability). In experiment 1 we show more accurate performance on a speeded name verification task for identities learned in high compared to low variability, when the test images are completely novel photos. In experiment 2 we show more accurate performance on a face matching task for identities previously learned in high compared to low variability. The results show that exposure to a large range of within-person variability leads to enhanced learning of new identities.

Introduction

It is well established that we are poor at recognising unfamiliar people. Matching different images of unfamiliar faces is a difficult task for human observers (Bruce et al, 1999, 2001; Clutterbuck & Johnston, 2002, 2004; Megreya & Burton 2006, 2007). Furthermore, in real-world settings people are highly error-prone when matching a live person to a photo-ID (Kemp, Towell & Pike, 1997). Experienced passport officers show poor levels of matching ability, no better than untrained students (White, Kemp, Jenkins, Matheson & Burton, 2014). In contrast to these poor performance levels for unfamiliar faces, recognition of familiar faces is excellent, remaining robust even in highly distorted or degraded images (Burton et al, 1999; Hole, George, Eaves & Rasek, 2002). Given this discrepancy between familiar and unfamiliar faces, it is imperative to understand how we learn new faces: how representations make the transition from unfamiliar to familiar. However, rather little is known about this process.

Much of the recent work on face learning has emphasised the systematic exposure to different photos of the same person, asking how learning-set images generalise to novel test images. Popular dimensions on which to manipulate learning sets include head angle and illumination (Longmore, Liu & Young, 2008; Liu, Bhuiyan, Ward & Sui, 2009). Liu and colleagues showed that extensive training with different images of the same face presented in multiple head angles led to accurate recognition of the same face presented in a different illumination. However, training with different illuminations did not produce high accuracy when the same face was presented in different head angles. The authors demonstrate that pose plays a more important role than illumination in forming generalizable representations.

Longmore and colleagues (2008) demonstrated in six different experiments that learning pictures is far easier than learning people. Participants learned identities presented in full-face and profile views and were tested on the same images as well as in novel views and lighting. Participants were more accurate at recognising the image they had learned than an image showing a different head angle. Furthermore, participants who had learned an identity from only full-face photographs could recognise the same identity presented in a three-quarter view, but not in a profile view. The complementary pattern of results was found for identities learned in a profile view. Performance in the three-quarter view condition was no better when participants learned two head angles than one.

The Longmore et al (2008) and Liu et al (2009) studies suggest that learning of unfamiliar faces, measured by subsequent recognition, is dependent on viewpoint. These studies are consistent with work on perceptual learning. Intensive training has been shown to improve fine-grain discrimination of object and pattern features, with improvements in the visual domain being specific to the trained dimensions in more difficult tasks, but with some evidence of generalisation in easier tasks (see Fahle, 2005). Training in one face view has been shown to lead to long-lasting increased sensitivity to that specific face orientation, but generalises across face sizes and face identities (Bi, Chen, Weng, He & Fang, 2010). Training in one face orientation has also been shown to lead to exemplar-specific improvements in face recognition (Hussain, Sekuler & Bennett, 2009). However, both these studies used one computer-generated face image per identity.

In contrast to the experiments described above, which use systematically-controlled facial images, there has recently been a rise in interest in the notion of ‘ambient images’ (Jenkins, White, van Montfort & Burton, 2011; Murphy, Ipser, Gaigg & Cook, 2015; Sutherland et al, 2013). These are images which vary in unsystematic ways, for example a set of images which would be generated by an internet search on the name of a celebrity. Such image sets pose considerable problems for the experimenter, because it is clear that viewers can recognize a wide range of photos a familiar face, varying across a very large range of dimensions: e.g. changes in the person due to expression, ageing, health; changes in the environment such as lighting and pose; changes in the camera such as focal length, perspective setting etc. Such variations are often not recoverable, i.e. we can recognize a picture of Barak Obama without knowing the focal length of the camera used – even though this will have a very large effect on the image (Burton, Schweinberger, Jenkins & Kaufmann, 2015).

While there is a natural inclination for scientists to eliminate variability due to sources which cannot easily be quantified, we have argued that *unsystematic* variability is, in fact, critical to face learning (Burton, 2013; Jenkins et al, 2011). This argument rests on the notion that individual faces have *idiosyncratic* variability. So, the ways in which one face varies are not (completely) the same as the ways in which another varies. If this is true, it explains the long-term puzzle of face matching. People find it hard to know whether two pictures of an unfamiliar person come from the range of images that could possibly depict that person, because they have no prior experience of this range. On the other hand, the wide experience with a familiar face provides one with a representation of the variability possible for that person – and hence it is easy to judge whether two novel images fall within it. In recent work, we have demonstrated that there is supporting evidence for this proposal in analysis of the physical properties of different images of the same person (Burton, Kramer, Ritchie & Jenkins, in press). People do, indeed, present variation in idiosyncratic ways, and so this information is available to support the human learning system.

One recent study has investigated the role of within-person variability on face learning by exposing participants to naturally occurring, or ambient images (Murphy et al, 2015). Participants saw the same eight identities on each trial, with either the same six images of each identity repeated on each trial, or with a new six images of each identity appearing on each trial The latter condition resulted in 96 unique images of each identity being viewed, and so participants in this condition were exposed to a wider range of variability within each identity. Exposure to variability within each identity led to a trial-by-trial decrease in estimation of the number of identities present, showing that participants were able, over time, to cohere together the images of each identity. Participants were also better at subsequently identifying the people for whom they had seen more images and variability, supporting previous work showing improved learning with increased numbers of exposures (see Shapiro & Penrod, 1986).

In the present studies, we manipulate within-person variability in a face learning task, while keeping the number of unique images in each condition consistent. We create high variability sets of images through simple internet search for foreign celebrities, unknown to our participants. However, we also create low-variability sets by taking individual frames from video interviews with these celebrities. Such low-variability images vary in pose and expression, but are similar in other respects (e.g., age, hairstyle, lighting, camera). Our hypothesis is that learning a new face through exposure to high variability images will support better learning (i.e. better recognition of a novel photo seen later) than exposure to low variability images.

Experiment 1a

In this experiment, we used a face/name association technique. In the learning phase participants were shown a series of photographs of each identity paired with that person's name. The test phase was a speeded name verification task in which participants were shown a name followed by a photograph, and asked to indicate whether or not the photograph showed the person named. In the learning phase, half of the identities were learned from ten highly variable photographs, the other half from ten less variable photographs. The images in the test phase were always completely novel.

Methods

Participants

Twenty participants (7 male; mean age: 22 years, range: 18-30 years) took part. All were students or other members of Aberdeen University.

Materials

Images of each of twenty Australian celebrities (ten female) were taken from a UK/Australian database, developed for bi-lateral research. Pre-checks with UK participants from our testing population (but not participating in the main experiment) confirmed that these Australian faces were not familiar to our participant group. The 'high variability' images were obtained from a Google Image search. Each image showed the full head, and was unconstrained in terms person variability such as facial expression, and in terms of environmental variability such as lighting and camera characteristics (see figure 1A). The 'low variability' images were stills from interview videos of each of the celebrities. These images varied in head angle and facial expression, but not other person-specific dimensions such as hair style or age. These 'low variability' images did not vary greatly in lighting or camera characteristics (see figure 1B). Background information was visible in all images. The images were presented centrally measuring 5.2 x x7.8cm on a 37.7 x 30.2 screen. The experiment was run using Matlab Pyschtoolbox.

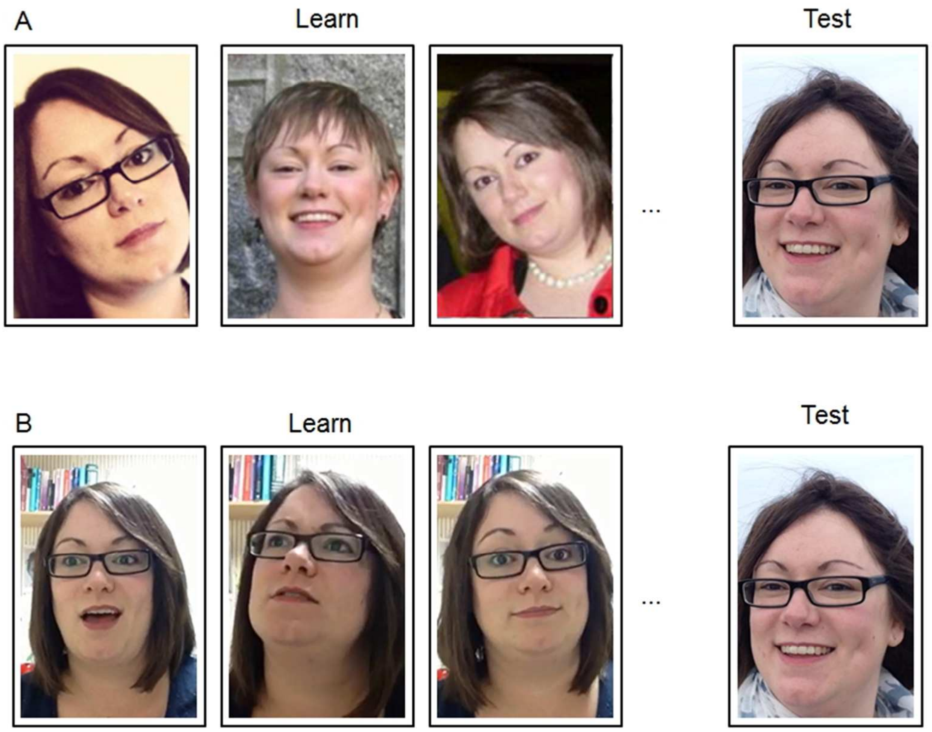

Figure 1. Example faces for experiment 1a. A, high variability images of the same person taken from a Google search (in the case of the experimental stimuli). B, low variability images of the same person taken from one interview video. The test image was always a previously unseen image. [Copyright restrictions prevent publication of the actual images used, though these are available from the authors. Images in figure1, also used in figure 3, are illustrative of the experimental stimuli, and depict someone who did not appear in the experiments but has given permission for the images to be reproduced here.]

Procedure

In the learning phase, participants saw ten images of each of the twenty faces, blocked by identity. Each image was presented centrally with the name of the person presented above. These were the real names of the Australian celebrities depicted and they remained on the screen throughout the block of images of that given ID. Individual images were presented for 5sec and there was an inter-stimulus interval of 500ms. Participants were instructed to try to learn each person, and there was a short rest break between blocks of faces. Participants saw the high variability set for half of the identities and the low variability set for the remaining identities, counterbalanced across identities between participants.

The test phase of the experiment was a speeded name verification task. A name was displayed on the screen for 1500ms, and immediately replaced by an image which remained on the screen until response. On half of the trials, the face matched the name. Participants responded via button press indicating whether the image showed the same person as the name or not, and were instructed to respond as quickly and as accurately as possible. The response buttons were counterbalanced between participants with half responding "same" with the left hand and half with the right. This phase used ten completely novel images of each identity (never seen during the learning phase). Match trials showed the person named, and mismatch trials showed an image of a different person of the same sex. There were five match trials and five mismatch trials for each identity.

Results and Discussion

None of the participants indicated prior familiarity with any of the identities. Figure 2 shows mean accuracy and reaction time data, with RTs based only on those trials on which participants responded correctly.

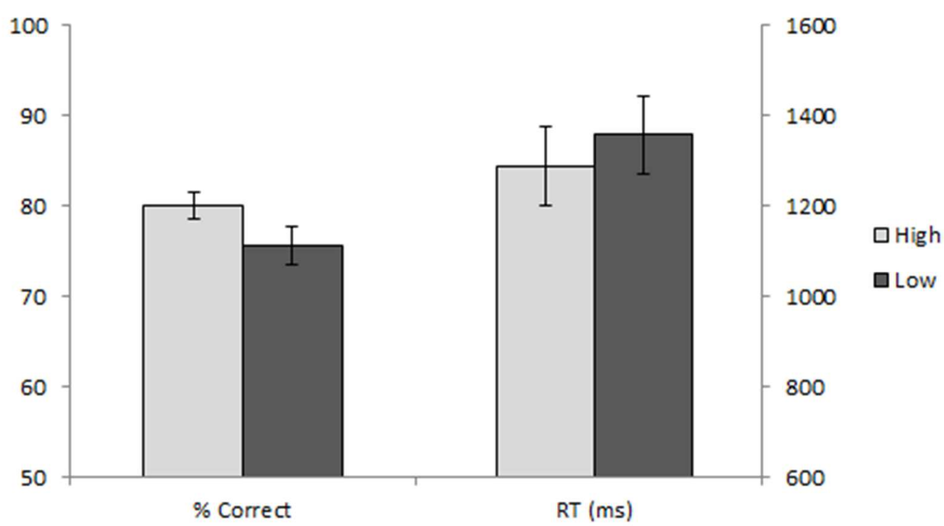

Figure 2. Accuracy and reaction time data for experiment 1a. Light bars denote identities learned from a high variability set, dark bars denote identities learned from a low variability set. Error bars denote the standard error of the mean.

A paired samples t-test on mean accuracy shows better performance for faces learned in high than low variability ($t(39) = 2.22$, $p < .05$). Reaction times are shown only for trials on which a correct response was given, and reveal no difference between conditions ($t(39) = 1.67$, $p > .05$).

In this experiment, participants were more accurate at responding to identities they had learned in high compared to low variability. The test images in the previous experiment were always completely novel, yet the possibility remains that this advantage arises because images at test more closely resemble some of the high than low variability learning images. To rule this out as a possible explanation, experiment 1b uses previously unseen low variability images at test.

Experiment 1b

This experiment is identical to the previous study, except that here we use previously unseen low variability images at test. These images by nature closely resemble the low variability images in the learning phase. If mere image similarity between learning and test is the explanation for our results in experiment 1a, then we would expect more accurate responses to images learned in low compared to high variability when tested on low variability images here.

Methods

Participants

A further twenty participants (7 male; mean age: 24 years, range: 19-61 years) took part, none of whom had participated in the previous experiment. All were students or other members of Aberdeen University.

Materials and Procedure

The materials and procedure were very similar to those used in experiment 1a. The images used in the learning phase were those used in experiment 1a with participants learning 20 IDs from high and low variability image sets. The test images were previously unseen images from the low variability set (see figure 3). All participants saw the same test images at test.

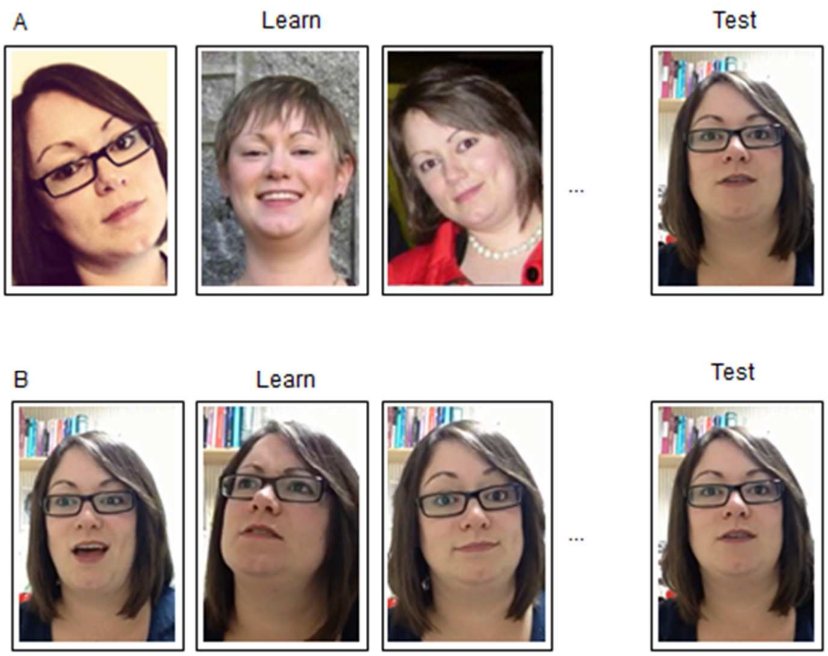

Figure 3. Example stimuli for experiment 2. A, high variability images taken from a Google search (in the case of the experimental stimuli). B, low variability images taken from one interview video. The test image was always a previously unseen low variability image.

Results and Discussion

Again, none of the participants indicated prior familiarity with any of the identities. Figure 4 shows mean accuracy and RT data.

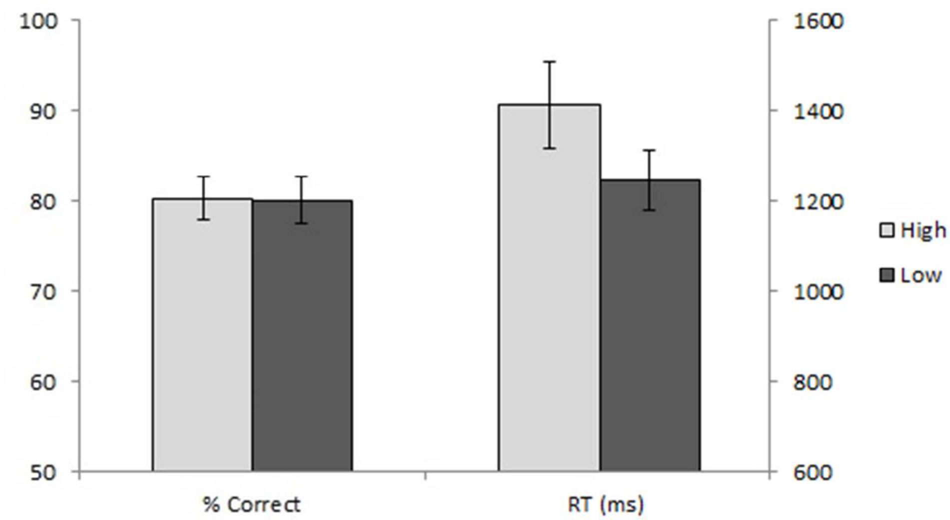

Figure 4. Accuracy and reaction time data for experiment 1b. Light bars denote identities learned from a high variability set, dark bars denote identities learned from a low variability set. Error bars denote the standard error of the mean.

A paired samples t-test shows no effect on accuracy, meaning that participants were no more accurate with identities they had learned from low or high variability image sets ($t(39) = 0.07$, $p > .05$, $\eta_p^2 = .00$). Again, reaction times are reported only for trials on which a correct response was given, and show faster responses for identities learned in low compared to high variability ($t(39) = 2.70$, $p < .05$, $\eta_p^2 = .16$).

As the test images in this experiment closely resembled the low variability learning images, the effect of variability on the speed of name verification is not surprising. Here we demonstrate an RT effect akin to image similarity effects found in perceptual learning experiments (Bi et al, 2010; Hussain et al, 2009; Longmore et al, 2008). What is surprising is that despite the dissimilarity between the test images and the high variability learning images, participants are able to identify the person pictured. This is a potentially surprising result as one might expect the similarity between test and learning images for the low variability set (including background information) to lead to a benefit for those faces. Instead learning people in high variability confers on viewers the ability to extrapolate well – to the extent here that performance is equivalent to having learned images similar to those used at test. Our RT effect shows that the process involved in making this decision takes longer than when the images at learning at test are very similar.

Experiment 2

This experiment employs a different dependent measure to assess the effect of variability on face learning. Here, participants learn identities through high or low variability and then complete a face matching task with new images of the learned identities, as well as new identities. As face matching tasks have been shown to be good predictors of familiarity (Clutterbuck & Johnston, 2002, 2004) we expect to see more accurate face matching for identities learned in high compared to low variability.

Methods

Participants

Thirty participants (4 male; mean age: 19.3 years, range: 18-22 years) took part. All were students at the University of York.

Materials

The stimuli for the learning phase were images for ten of the faces used in experiment 1 (5 female). For the matching phase of the experiment, completely novel images were used, two of each identity for match trials, and one image of the target identity and one foil image for mismatch trials. Foil images were obtained from a Google image search for the verbal description of each ID. Images of 80 new identities were also used in the matching phase to provide a baseline of matching performance for identities which had not been learned.

Procedure

The learning phase followed the procedure used in experiment 1. Participants saw the high variability set for half of the identities and the low variability set for the remaining identities, counterbalanced across identities between participants.

The test phase of the experiment was a face matching task. Two images were displayed on the screen and participants indicated via button press whether the two images showed the same person or two different people. The images remained on screen until a response was made. Half of the trials were match trials and half mismatch. There were eight match and eight mismatch trials for each of the learned identities, using novel images of each identity not seen in the learning phase, and one match and mismatch trials for each of the new identities. The new identities were simply shown so as to provide a measure of baseline face matching performance for identities which had not been previously learned.

Results and Discussion

None of the participants indicated prior familiarity with any of the identities. Figure 5 show mean matching performance across conditions.

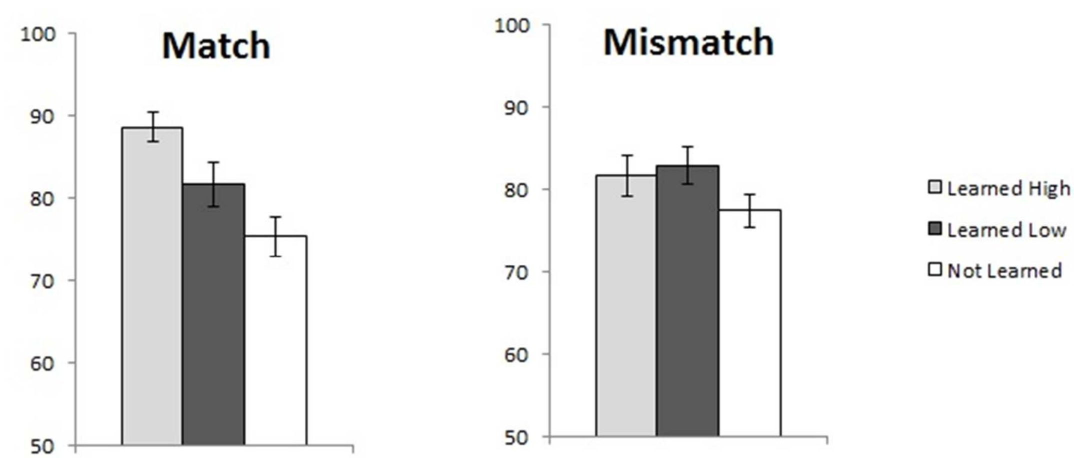

Figure 5. Accuracy on face matching task. Light bars denote identities learned from a high variability set, dark bars denote identities learned from a low variability set, white bars denote identities that were not previously learned. Error bars denote the standard error of the mean.

We analysed match and mismatch trials separately as in previous studies (e.g., White, Burton, Jenkins & Kemp, 2014). A one-way ANOVA on match trials showed a main effect of condition ($F(2,58) = 11.29$, $p < .001$, $\eta_p^2 = .28$), with follow-up Tukey HSD comparisons showing higher accuracy for IDs learned in high compared to low variability, and those learned in high variability compared to those not learned (both $p < .05$). The difference between accuracy for IDs learned in low variability compared to those not learned was not significant ($p > .05$). ANOVA on mismatch trials showed no effect of condition ($F(2,58) = 2.05$, $p > .05$).

The results of this experiment show that learning a person through a highly variable set of images leads to improved face matching. Interestingly, the effect is present only in match trials and not mismatch trials. Although some previous work has shown that familiarity can improve performance on both match and mismatch trials (Clutterbuck & Johnston, 2002, 2004, 2005), our pattern of results is not unprecedented. In many previous studies investigating face matching, the experimental manipulation resulted in a change in either match or mismatch performance, but rarely both (e.g. Megreya & Burton, 2006, 2007; White et al, 2014; Menon, White & Kemp, 2015). In a series of experiments investigating the comparative benefits of average images and arrays of images for face matching tasks, White et al (2014) found advantages only for match trials. In fact, previous work has shown that accuracy on match and mismatch trials is not correlated (Megreya & Burton, 2007) and so our dissociation between match and mismatch performance could be indicative of two different processes at work during face learning. Through our learning procedure, participants learn to incorporate and tolerate high levels of within-person variability, without this leading to a general acceptance of all images including images of foil identities. It is worth noting that the same pattern of results exists in experiments 1a and 1b whereby the general performance increase observed is due to elevated performance on trials where the name and face depict the same person with no differences on trials where they depict different people.

General Discussion

Our results show that experiencing a relatively wide range of within-person variability from face photographs produces better learning than experiencing a narrower range of variability. Furthermore, we have demonstrated this with two rather different measures. Exposure to high variability images leads to better name verification (experiment 1a) and better matching (experiment 2) as compared to exposure to low variability images. Furthermore, even when task demands favor learning of image-specific properties (experiment 1b), exposure to highly variable learning images is very effective.

In our experiments, participants saw images of each ID sequentially as they learned the new identity. The results suggest that participants were able to build up a representation of each identity over time during this sequence. This supports a recent study which has shown that participants build up a representation of someone’s appearance over time, with identity manipulations influencing performance on a sequential but not simultaneous matching task

(Menon, White & Kemp, 2015). We suggest that experiencing variability over time supports an abstraction of representation of identity and may be a route to face learning.

In all of the studies presented here we have compared learning from high variability to learning from low variability, yet we should point out that even our "low variability" image set incorporates a wider range of variability than is traditionally seen in face learning experiments. Our use of ambient images taken from internet searches and screen captures of videos is not simply an attempt to appeal to a sense of "ecological validity" or, indeed, to suggest that variability seen in the real world is confined to that seen in our experiment. We are simply seeking to use less controlled stimuli than have been used in the past in order to preserve within-person variability. Even our low variability stimuli incorporate, in an uncontrolled way, changes in head angle and lighting which have been studied under controlled conditions previously (Liu et al, 2009; Longmore et al, 2008). Our low variability stimuli also show variance in expression and small changes in distance from camera. Our high variability sets include all of these forms of within-person variance, as well as changes in weight, age, hairstyle, makeup etc. This shows a naturally occurring, wide range of differences within each individual identity. We have argued previously that much of this variability is idiosyncratic, and is critical to a perceiver becoming familiar with a specific person (Burton, 2013; Burton, Jenkins & Schweinberger, 2011; Burton et al, in press; Jenkins et al, 2011). For this reason it seems likely that even our low variability stimuli give rise to more accurate face learning than repeated, or extended, presentation of a single image. Previous studies using single-image learning have shown better learning for longer exposure durations. (e.g. MacLin, MacLin & Malpass, 2001; Read, Vokey & Hammersley, 1990; Reynolds & Pezdek, 1992). It is therefore possible that our low-variability images mimic repeated presentation of single images in viewers' perceptions, though we think it unlikely (see figure 3A). How much variability is necessary to benefit learning, or to optimise it, is an interesting question for future research.

We did not employ a training technique in our learning experiments as has been used previously to ensure 100% accurate face learning prior to test (Longmore et al, 2008). The use of such a technique ensures that participants have learned each of the initial images of each identity accurately before testing takes place with a new set of images. Crucial to our design is keeping number of exposures consistent across identities and conditions, something not always achieved through the use of training techniques. Furthermore, we did not want participants to learn specific images of the identities, rather we were allowing participants to construct their own mental representation of each of these new people in order to try to learn them – an approach which, we argue, more naturally mimics real life face learning.

Many authors have proposed that recognition of unfamiliar faces is predominantly image-bound (Bruce, 1982; Hancock, Burton & Bruce, 2000; Megreya & Burton, 2008) and so in traditional face learning paradigms varying pose and lighting, the similarity between the images presented at learning and test has been a good predictor of performance (Braje, 2003; Krouse, 1981; Longmore et al, 2008). However, such studies eliminate the possibility of studying ambient variation in the images of faces that people actually learn in daily life. It seems that this variation is, in fact, a very powerful source of information for people acquiring representations of novel faces. If one is able to extract information about the idiosyncratic ways in which people vary, then this can lead to generalizable representation, useful across many types of novel visual context (Burton, 2013; Jenkins et al, 2011).

In addition to informing our understanding of face learning, these results also have interesting practical implications. Previous studies investigating the ability of security professionals to recognise unfamiliar faces has shown that relatively poor performance by passport officers (White et al, 2014), and police (Krouse, 1981). The ability to familiarise viewers with new identities relatively quickly through exposure to within-person variability may be useful in applied settings. For example, passport-issuing staff who are required to recognise known criminals applying for new passports under false identities, or police looking for a particular suspect in a crowd may be able to exploit this fast route to learning a new identity for practical purposes.

Acknowledgements

The research leading to these results has received funding from the European Research Council under the European Union’s Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n.323262, and from the Economic and Social Research Council, UK [ES/J022950/1].

References

Bi, T., Chen, N., Wneg, Q., He, D., & Fang, F. (2010). Learning to discriminate face views. *Journal of Neurophysiology, 104,* 3305-3311.

Braje, W. L. (2003). Illumination encoding in face recognition: Effect of position shift. *Journal of Vision, 3,* 161–170.

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

Bruce, V., Henderson, Z., Greenwood, K., Hancock, P. J. B., Burton, A. M., & Miller, P. (1999). Verification of identities from images captured on video. *Journal of Experimental Psychology: Applied, 5(4),* 339-360.

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.

Burton, A. M. (2013). Why has research in face recognition progressed so slowly? The importance of variability. *Quarterly Journal of Experimental Psychology*, *66(8)*,1467-1485.

Burton, A. M., Jenkins, R., & Schweinberger, S. R. (2011). Mental representations of familiar faces. *British Journal of Psychology, 102,* 943-958.

Burton, A. M., Kramer, R. S. S., Ritchie, K. L., & Jenkins, R. (in press). Identity from variation: Representations of faces derived from multiple instances. *Cognitive Science.*

Burton, A. M., Schweinberger, S., Jenkins, R., & Kaufmann, J. (2015). Arguments against a configural processing account of familiar face recognition. *Perspectives on Psychological Science, 10(4),* 482-496.

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.

Clutterbuck, R., & Johnston, R. A. (2002). Exploring levels of face familiarity by using an indirect face-matching measure. *Perception, 31,* 985–994.

Clutterbuck, R., & Johnston, R. A. (2004). Matching as an index of face familiarity. *Visual Cognition, 11(7),* 857-869.

Clutterbuck, R., & Johnston, R. A. (2005). Demonstrating how unfamiliar faces become familiar using a face matching task. *European Journal of Cognitive Psychology, 17(1),* 97-116.

Fahle, M. (2005). Perceptual learning: Specificity versus generalization. *Current Opinion in*

Neurobiology, 15, 154-160.

Hancock, P. J. B., Bruce, V., & Burton, A. M. (2000). Recognition of unfamiliar faces. *Trends in Cognitive Sciences*, *4*(9), 330–337.

Hole, G. J., George, P. A., Eaves, K., & Rasek, A. (2002). Effects of geometric distortions on face-recognition performance. *Perception, 31,* 1221-1240.

Hussain, Z., Sekuler, A. B., & Bennett, P.J. (2009). Perceptual learning modifies inversion effects for faces and textures. *Vision Research, 49,* 2273-2284.

Jenkins, R., White, D., van Montfort, X., & Burton, A. M. (2011). Variability in photos of the same face. *Cognition, 121,* 313-323.

Kemp, R., Towell, N., & Pike, G. (1997). When seeing should not be believing: Photographs, credit cards and fraud. *Applied Cognitive Psychology*, *11*(3), 211–222.

Krouse, F. L. (1981). Effects of pose, pose change, and delay on face recognition performance. *Journal of Applied Psychology, 66,* 651–654.

Liu, C. H., Bhuiyan, A., Ward, J., & Sui, J. (2009). Transfer between pose and illumination training in face recognition. *Journal of Experimental Psychology: Human Perception and Performance, 35(4),* 939-947.

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

MacLin, O. H., MacLin, M. K., & Malpass, R. S. (2001). Race, arousal, attention, exposure, and delay: An examination of factors moderating face recognition. *Psychology, Public Policy, and Law, 7(1),* 134-152.

Megreya, A. M., & Burton, A. M. (2006). Unfamiliar faces are not faces: Evidence from a matching task. *Memory & Cognition*, *34*(4), 865–876.

Megreya, A. M., & Burton, A. M. (2007). Hits and false positives in face matching: A familiarity-based dissociation. *Perception & Psychophysics*, *69*(7), 1175.

Megreya, A. M., & Burton, A. M. (2008). Matching faces to photographs: Poor performance in eyewitness memory (without the memory). *Journal of Experimental Psychology: Applied, 14(4),* 364-372.

Menon, N., White, D., & Kemp, R. I. (2015). Identity-level representations affect unfamiliar face matching performance in sequential but not simultaneous tasks. *The Quarterly Journal of Experimental Psychology.* Advance online publication.

Murphy, J., Ipser, A., Gaigg, S., & Cook, R. (2015). Exemplar variance supports robust learning of facial identity. *Journal of Experimental Psychology: Human Perception and Performance.* Advance online publication.

Read, J. D., Vokey, J. R., & Hammersley, R. (1990). Changing photos of faces: Effects of exposure duration and photo similarity on recognition and the accuracy-confidence relationship. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16(5),* 870-882.

Reynolds, J. K., & Pezdek, K. (1992). Face recognition memory: The effects of exposure duration and encoding instruction. *Applied Cognitive Psychology, 6,* 279-292.

Shaprio, P. N., & Penrod, S. (1986). Meta-analysis of facial identification studies. *Psychological Bulletin, 100,* 139-256.

Sutherland, C. A. M., Oldmeadow, J. A., Santos, I. M., Towler, J., Burt, D. M., & Young, A. W. (2013). Social inferences form faces: Ambient images generate a three-dimensional model. *Cognition, 127(1),* 105-118.

White, D., Burton, A. M., Jenkins, R., & Kemp, R. I. (2014). Redesigning photo-ID to improve unfamiliar face matching performance. *Journal of Experimental Psychology: Applied, 20(2)*, 166-173.

White, D., Kemp, R. I., Jenkins, R., Matheson, M., & Burton, A. M. (2014). Passport Officers' Errors in Face Matching. *PLoS One, 9*, 8, e103510.