City, University of London Institutional Repository

Exemplar variance supports robust learning of facial identity

Jennifer Murphy, Alberta Ipser, Sebastian Gaigg, & Richard Cook*

Department of Psychology, City University London, London, U.K.

*Corresponding author: Richard.Cook.1@city.ac.uk

Department of Psychology,

City University London,

Whiskin Street

London, EC1R OJD

Tel: +44 20 7040 8644, Fax: +44 20 7040 8877

Abstract

Differences in the visual processing of familiar and unfamiliar faces have prompted considerable interest in face learning, the process by which unfamiliar faces become familiar. Previous work indicates that face learning is determined in part by exposure duration; unsurprisingly, viewing faces for longer affords superior performance on subsequent recognition tests. However, there has been further speculation that exemplar variation, experience of different exemplars of the same facial identity, contributes to face learning independently of viewing time. Several leading accounts of face learning, including the averaging and pictorial coding models, predict an exemplar variation advantage. Nevertheless, the exemplar variation hypothesis currently lacks empirical support. The present study therefore sought to test this prediction by comparing the effects of unique exemplar face learning – a condition rich in exemplar variation – and repeated exemplar face learning – a condition that equates viewing time, but constrains exemplar variation. Crucially, observers who received unique exemplar learning displayed better recognition of novel exemplars of the learned identities at test, than observers in the repeated exemplar condition. These results have important theoretical and substantive implications for models of face learning and for approaches to face training in applied contexts.

Key words:

Face learning, familiarity, face recognition, internal feature advantage, exemplar variation, pictorial coding, averaging.

Introduction

Familiar and unfamiliar faces engage different types of visual processing (Burton & Jenkins, 2011; Hancock, Bruce, & Burton, 2000; Jenkins & Burton, 2011; Megreya & Burton, 2006). As faces become more familiar, observers are better able to match targets using their internal features (eyes, nose, mouth), the so-called 'internal feature advantage'. In contrast, unfamiliar face matching is frequently based on external features, such as hairstyle and face shape (Ellis, Shepherd, & Davies, 1979; Osborne & Stevenage, 2008; Young, Hay, McWeeny, Flude, & Ellis, 1985). Familiar faces may also place lower demands on visual working memory (Jackson & Raymond, 2008) and are easier to detect under conditions of reduced attention (Jackson & Raymond, 2006), compared with unfamiliar faces. However, perhaps the most striking difference between familiar and unfamiliar face perception is the ease with which we can recognise new exemplars. In contrast to the effortless recognition of celebrities, colleagues and friends, matching the faces of strangers across different photographic images can be remarkably difficult (Bruce et al., 1999; White, Kemp, Jenkins, Matheson, & Burton, 2014). For example, when asked to sort photographs of two individuals according to the identity of those depicted, observers perform poorly, frequently attributing the photographs to eight or more different individuals (Jenkins, White, Van Montfort, & Burton, 2011).

Differences in the visual processing of familiar and unfamiliar faces have prompted considerable interest in face learning, the process by which unfamiliar faces become familiar. Previous evidence suggests that face learning is determined, at least in part, by the time observers spend viewing faces. Unsurprisingly, participants allowed to observe faces for 45 secs each outperform those who view the same faces for 15 secs (Memon, Hope, & Bull, 2003). Similarly, simple repetition of single facial images can improve subsequent recognition of actors in dynamic video stimuli (Roark, O'Toole, Abdi, & Barrett, 2006) and increase the strength of identity adaptation (Jiang, Blanz, & O'Toole, 2007), thought to be causally related to recognition ability (e.g., Dennett, McKone, Edwards, & Susilo, 2012). Crucially however, there has been further speculation that exemplar variation – experience of different exemplars of the same facial identity, such as that provided by a series of photographs – contributes to face learning, independently of viewing time.

Several leading accounts of face learning predict an exemplar variation advantage. For example, according to the pictorial coding model, familiar faces are recognised through comparison with previously stored instances of that face (e.g. Longmore, Liu, & Young, 2008). Having encountered multiple exemplars of a given face, observers are able to densely sample the potential instance space. Thereafter, the likelihood of a close match between a novel target exemplar and a previously stored instance is high, yielding superior face recognition performance. In contrast, an observer who has previously encountered only a few exemplars of a facial identity must rely on a sparse sampling of the instance space, and may therefore struggle to match the target to a previously stored instance.

A second closely-related model is the averaging hypothesis (Benson & Perrett, 1993; Burton, Jenkins, Hancock, & White, 2005; Jenkins & Burton, 2011). According to this view, exemplar variation allows the visual system to form a robust average of that facial identity. Transient differences including variation in lighting, shadow, hairstyle, expression are discounted, leaving a stable representation of permanent, reliable features. Crucially, i) accurate recognition of the target face upon subsequent encounters is thought to be directly related to the quality of the average formed (e.g. Jenkins & Burton, 2008), and ii) average estimates derived from many observations are likely to better approximate the true parameter value than estimates derived from only a handful of observations of comparable quality.

Despite its theoretical significance, however, the exemplar variation hypothesis currently lacks empirical support. Specifically, there is little evidence that exemplar rich experience supports better face learning, when viewing time is equated. In the absence of empirical evidence in favour of the exemplar variation hypothesis, it is possible that the strength of face learning is determined solely by viewing time[1]. The present study therefore adopted a two stage training-test procedure to determine whether exposure to many unique exemplars of a target face ('unique exemplar learning') supports superior recognition, compared with equivalent exposure to a limited number of exemplars ('repeated exemplar learning'). Similar recognition performance following both types of training would indicate that face learning is determined not by exemplar variation, but by the time spent viewing the trained identities, challenging a key assumption of the pictorial coding and averaging accounts.

Methods

Participants

Fifty healthy adults (19 males, M_{age} = 29.5 years, SD_{age} = 8.1 years) participated in this experiment in return for a small honorarium. Participants were randomly allocated to the different training conditions in equal numbers. Four participants, two in each condition, were replacements. Two of the replaced observers scored at chance level at test, one reported prior familiarity with a learned identity, and one was an outlier in terms of response latencies. All participants had normal or corrected-to-normal vision, gave informed consent, and were fully debriefed upon task completion. Sample size was determined a-priori using power analysis assuming a large effect size (Cohen, 1988). Ethical clearance was granted by the local ethics committee and the study was conducted in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki.

Training

Participants completed one of two different face learning procedures, each comprising 16 trials. Each training trial presented 48 facial images simultaneously in a 6 × 8 array. Training arrays were presented for 48 seconds, during which observers were free to inspect the images as they wished. Following array offset, a prompt appeared to judge the number of identities represented within the array, as accurately as possible. Unbeknown to observers, the 48 images were in fact always taken from only eight individuals, training arrays comprising six images of each. The same eight individuals – the to-be-learned identities – were shown on every training trial. The training faces sampled a broad range of poses, expressions, hairstyles, lighting conditions, and camera parameters (Figure 1).

Observers in the unique exemplar condition saw a novel set of 48 images on each training trial. Across the procedure, they were therefore exposed to 96 photos of each of the eight to-be-learned identities, six novel exemplars of each identity on each trial. Observers in this condition viewed different combinations of exemplars, chosen at random by the program, on each of the sixteen training trials. Observers in the repeated exemplar condition however, were exposed to the same 48 images on all sixteen training trials. Consequently, observers saw only six photos of each of the eight to-be-

learned identities. Each observer in this condition was trained on a different set of exemplars, chosen at random from the 96 images comprising each of the eight identity sets. That a novel combination of 48 images was selected for each participant guarded against the possibility of systematic bias. While the same 48 images were repeated, the position of the exemplars in the 6 × 8 array was randomized across the 16 learning trials. In all other respects, the two learning conditions were identical.

Figure-1

Training images were cropped to square aspect ratios subtending 3° vertically when viewed in the array at a distance of 60 cm. The to-be-learned identity was the only face visible in each image and was central and prominent. In all other respects the raw photographic images were left unaltered. Images showing the to-be-learned identities wearing glasses or sunglasses were intentionally avoided. The next trial commenced only once a response had been recorded. Four practice trials incorporating cartoon faces from *The Simpsons* preceded the learning procedure. All experimental programs were written in MATLAB using Psychtoolbox (Brainard, 1997; Pelli, 1997), and presented on a Dell 17 inch LCD display at 60-Hz refresh rate.

Test

Immediately after training, all participants completed the same test procedure to assess their recognition of the eight learned identities. Unlike the training faces, test faces were presented in greyscale and were cropped to exclude external features, thereby removing the cues typically used to recognise unfamiliar faces (Ellis et al., 1979; Osborne & Stevenage, 2008; Young et al., 1985). Trials briefly (for 1000 ms) presented a single face centrally, followed by a prompt to judge whether the identity was or was not encountered during the learning phase ('old' or 'new'). Half of the facial images presented during test depicted the eight learned identities, however none of the exemplars were used during the training phase. The remaining test images were exemplars of eight novel identities not encountered previously. Test faces were presented in greyscale and subtended 6.5° vertically when viewed at 60 cm. All test faces were shown in frontal view and had approximately neutral expressions. The 'old' and 'new' test images were closely matched, both in terms of the identities depicted and

their presentation. In total the test procedure comprised 160 trials: Five novel exemplars of the eight trained identities and five exemplars of eight novel identities were presented twice each. The test was preceded by eight practice trials incorporating cartoon faces from *The Simpsons*.

Results

The identity estimates from the training phase (Figure 2) were analysed using ANOVA with trial (1:16) as a within-subjects factor and learning condition (unique exemplar, repeated exemplar) as a between-subjects factor. The analysis revealed highly significant linear [$F(1,48) = 37.250$, $p <.001$, $\eta^2 = .437$] and quadratic trends [$F(1,48) = 15.065$, $p <.001$, $\eta^2 = .239$], indicative of learning across trials. Neither the linear [$F(1,48) = .249$, $p = .620$, $\eta^2 = .005$] nor quadratic [$F(1,48) = .224$, $p = .638$, $\eta^2 = .005$] trends varied as a function of learning condition, indicating that the change in identity estimates across the training procedure was broadly comparable for the two groups. The main effect of learning condition was not significant [$F(1,48) = .558$, $p = .459$, $\eta^2 = .011$], suggestive of similar identity estimates overall. The mean identity estimates did not differ between the learning conditions on the first [$t(48) = .144$, $p = .886$] or last trials [$t(48) = .355$, $p = .724$]. Despite the learning observed, the identity estimates of both the unique exemplar [$t(24) = 16.249$, $p < .001$] and repeated exemplar [$t(24) = 13.226$, $p < .001$] groups still exceeded the true number (eight identities) on the final training trial.

Figure-2

The crucial differences between the groups were seen at test. As predicted by the exemplar variation hypothesis, the observers who received unique exemplar learning ($M = 80.8\%$, $SD = 10.2\%$) outperformed those who received repeated exemplar learning ($M = 72.9\%$, $SD = 8.0\%$) [$t(48) = 3.044$, $p = .004$]. This difference corresponds to a Cohen's d of 0.80, indicative of a large effect (Cohen, 1988). The advantage of unique exemplar learning ($M = 77.9\%$, $SD = 14.7\%$) over repeated exemplar learning ($M = 66.0\%$, $SD = 13.6\%$), was only present for hits; i.e., correct 'old' responses in the presence of an 'old' stimulus [$t(48) = 2.993$, $p = .004$]. No advantage of unique exemplar learning ($M = 83.7\%$, $SD = 11.1\%$) over repeated exemplar learning ($M =$

79.9%, *SD* = 18.0%) was observed for correct rejections; i.e., 'new' responses in the presence of a 'new' stimulus [$t(39.841) = .908$, $p = .369$ (corrected for inequality of variance)]. Of note, simple correlation analyses revealed that the identity estimates from the final training trial, an index of the performance achieved by each observer on completion of the training task, correlated significantly with test performance for the unique exemplar condition ($r = -.503$, $p = .010$), but not for the repeated exemplar condition ($r = .013$, $p = .952$).

Discussion

The present study compared the effects of unique exemplar face learning – a condition rich in exemplar variation – and repeated exemplar face learning – a condition that equates viewing time, but constrains exemplar variation. During training, observers in both learning conditions overestimated the number of individuals present, particularly at the start of the procedure, when the identities were novel. As the faces became more familiar, both groups found it easier to recognise the commonalities across different exemplars and their identity estimates became progressively more accurate. Crucially however, observers who received unique exemplar learning displayed better recognition of novel exemplars of the learned identities at test, than observers in the repeated exemplar condition. These results have important theoretical and substantive implications for models of face learning and for approaches to face training in applied contexts. Crucially, these results confirm for the first time that face learning is determined not only by the time spent viewing a face, but also the degree of exemplar variation experienced by the learner.

The finding that observers over-estimate the number of identities present in arrays comprising unfamiliar faces is consistent with previous reports (Jenkins et al., 2011), and further underscores the challenges posed by unfamiliar face matching (Bruce et al., 1999; White et al., 2014). However, in both learning conditions identity estimates declined sharply across the first 8-10 training trials, revealing that a degree of familiarity can be acquired rapidly based on relatively little exemplar variation. Interestingly, the identity estimates from the final training trial, an index of the performance achieved by each observer on completion of the training task, correlated significantly with test performance for the unique exemplar condition, but not for the

repeated exemplar condition. This unexpected finding raises the possibility that the nature of the leaning observed in the two conditions was qualitatively different. In light of previous findings (e.g., Osborne & Stevenage, 2008), we speculate that observers in the repeated exemplar condition were disproportionately reliant on the external facial features during the training procedure, whereas those allocated to the unique exemplar condition grew more adept at distinguishing faces from their internal features. When forced to discriminate faces based solely on internal features at test, performance therefore correlated with individual differences in unique exemplar learning, but not with the variability seen in repeated exemplar learning.

The present results have interesting implications for the study of moving faces. When considered as a time-series of different facial poses and expressions, it is clear that even a short sequence of facial motion contains many unique exemplars. Viewing moving faces might therefore be expected to convey a face learning advantage relative to the observation of a few static images (O'Toole, Roark, & Abdi, 2002). Paradoxically however, many studies have found little or no evidence for such a motion advantage when learning new faces (e.g., Bonner, Burton, & Bruce, 2003). One possibility is that the rigid and non-rigid movements present within naturalistic sequences of facial motion may be relatively constrained, and therefore fail to yield sufficient exemplar variation to generate a clear learning advantage.

Overall, the present findings show for the first time that face learning is determined not only by the time spent viewing a face, but also the degree of exemplar variation experienced by the learner. That unique exemplar learning affords recognition performance superior to repeated exemplar learning, confirms that exemplar variation contributes to face learning in human vision. These results suggest that face training applications should seek to maximise exemplar variation to optimise face learning.
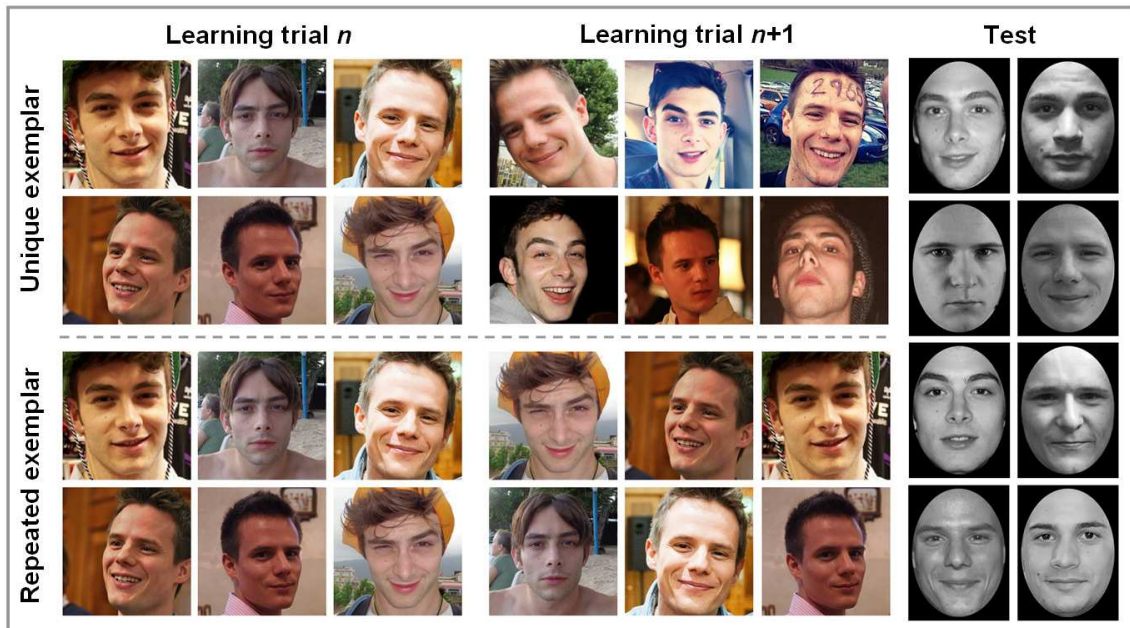
Figure 1

Figure 1: Examples of the training images and illustration of the learning manipulation (left). Observers in the unique exemplar learning condition were exposed to six new photos of each learned identity on each training trial. In the repeated exemplar condition, observers were exposed to the same six photos on every learning trial. Examples of the test images (right).
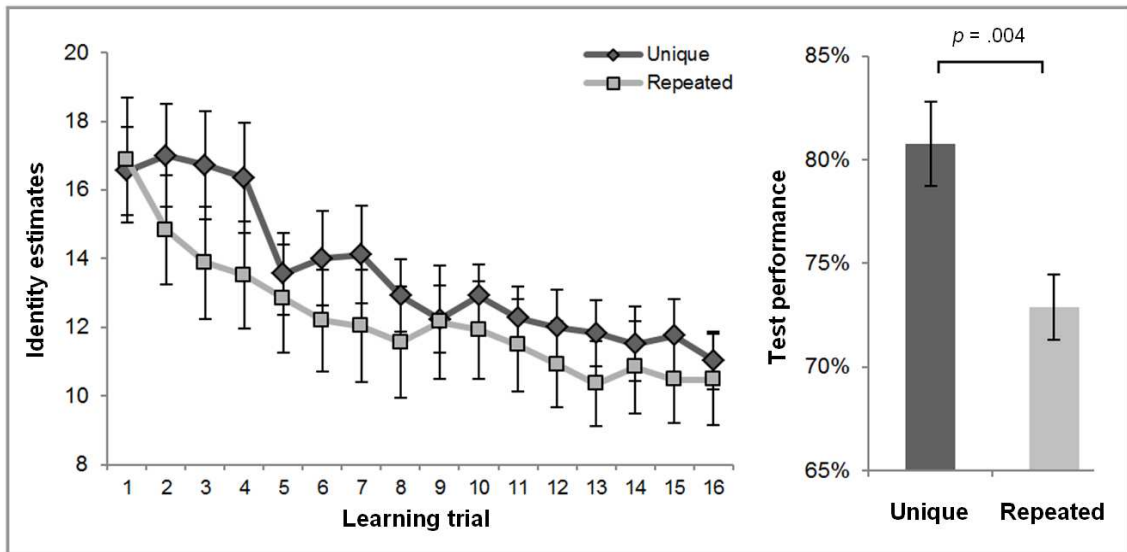
Figure 2

Figure 2: Results from the training phase (left). Observers in both conditions overestimated the number of individuals present in the training arrays, particularly at the start of the procedure. Identity estimates declined steadily thereafter indicative of learning. Results from the test phase (right). Participants were briefly presented cropped greyscale facial images, and asked whether or not the identity depicted was encountered during the learning phase ('old' or 'new'). In accordance with the exemplar variation hypothesis, unique exemplar learning supported better recognition than repeated exemplar learning. Error bars indicate ± one standard error of the mean.

Footnote:

[1]For example, having experienced how familiar faces appear in different poses and lighting situations during the course of development, adults may be able to extrapolate from, or interpolate between, a handful of exemplars of a novel target face. Should observers be able to generate variation endogenously, through the application of previously learned image transformations, the information conveyed by additional unique exemplars may be redundant.

Acknowledgements

References

Benson, P. J., & Perrett, D. I. (1993). Extracting prototypical facial images from exemplars. *Perception, 22*(3), 257-262.

Bonner, L., Burton, A. M., & Bruce, V. (2003). Getting to know you: How we learn new faces. *Visual Cognition, 10*(5), 527–536.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*(4), 433-436.

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

Burton, A. M., & Jenkins, R. (2011). Unfamiliar face perception. In A. J. Calder, G. Rhodes, J. V. Haxby & M. H. Johnson (Eds.), *The Oxford handbook of face perception* Oxford, UK: Oxford University Press.

Burton, A. M., Jenkins, R., Hancock, P. J., & White, D. (2005). Robust representations for face recognition: the power of averages. *Cognitive Psychology, 51*(3), 256-284.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences (Second Edition)*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Dennett, H., McKone, E., Edwards, M., & Susilo, T. (2012). Face aftereffects predict individual differences in face recognition ability. *Psychological Science, 23*(11), 1279-1287.

Ellis, H. D., Shepherd, J. W., & Davies, G. M. (1979). Identification of familiar and unfamiliar faces from internal and external features: some implications for theories of face recognition. *Perception, 8*(4), 431-439.

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

Jackson, M. C., & Raymond, J. E. (2006). The role of attention and familiarity in face identification. *Perception and Psychophysics, 68*(4), 543-557.

Jackson, M. C., & Raymond, J. E. (2008). Familiarity enhances visual working memory for faces. *Journal of Experimental Psychology: Human Perception and Performance, 34*(3), 556-568.

Jenkins, R., & Burton, A. M. (2008). 100% accuracy in automatic face recognition. *Science, 319*(5862), 435.

Jenkins, R., & Burton, A. M. (2011). Stable face representations. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 366*(1571), 1671-1683.

Jenkins, R., White, D., Van Montfort, X., & Burton, A. M. (2011). Variability in photos of the same face. *Cognition, 121*(3), 313-323.

Jiang, F., Blanz, V., & O'Toole, A. J. (2007). The role of familiarity in three-dimensional view-transferability of face identity adaptation. *Vision Research, 47*(4), 525-531.

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.

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

Memon, A., Hope, L., & Bull, R. (2003). Exposure duration: effects on eyewitness accuracy and confidence. *British Journal of Psychology, 94*(3), 339-354.

Osborne, C. D., & Stevenage, S. V. (2008). Internal feature saliency as a marker of familiarity and configural processing. *Visual Cognition, 16*(1), 23-43.

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

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision, 10*(4), 437-442.

Roark, D. A., O'Toole, A. J., Abdi, H., & Barrett, S. E. (2006). Learning the moves: The effect of familiarity and facial motion on person recognition across large changes in viewing format. *Perception*, *35*(6), 761-773.

White, D., Kemp, R. I., Jenkins, R., Matheson, M., & Burton, A. M. (2014). Passport officers' errors in face matching. *PLoS One, 9*(8), e103510.

Young, A. W., Hay, D. C., McWeeny, K. H., Flude, B. M., & Ellis, A. W. (1985). Matching familiar and unfamiliar faces on internal and external features. *Perception, 14*(6), 737-746.