

Face-Space Architectures: Evidence for the Use of Independent Color-Based Features

Psychological Science
24(7) 1294–1300
© The Author(s) 2013
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0956797612464889
pss.sagepub.com


Adrian Nestor, David C. Plaut, and Marlene Behrmann

Center for the Neural Basis of Cognition and Department of Psychology, Carnegie Mellon University

Abstract

The concept of psychological face space lies at the core of many theories of face recognition and representation. To date, much of the understanding of face space has been based on principal component analysis (PCA); the structure of the psychological space is thought to reflect some important aspects of a physical face space characterized by PCA applications to face images. In the present experiments, we investigated alternative accounts of face space and found that independent component analysis provided the best fit to human judgments of face similarity and identification. Thus, our results challenge an influential approach to the study of human face space and provide evidence for the role of statistically independent features in face encoding. In addition, our findings support the use of color information in the representation of facial identity, and we thus argue for the inclusion of such information in theoretical and computational constructs of face space.

Keywords

color perception, face recognition, independent component analysis, principal component analysis, face perception, facial features, computer simulation, face space

Received 5/21/12; Revision accepted 9/23/12

The concept of face space (Valentine, 1991) has been seminal in the study of human face recognition, providing a theoretical framework for extensive empirical work and connecting this work with the computational study of face recognition (O’Toole, 2011). In essence, psychological face space refers to an internal multidimensional space in which individual faces can be represented as single points. This space enables the description of individual faces with respect to a limited number of perceptual dimensions and with respect to each other in terms of space-based distances.

Much research has focused on characterizing the structure and the dimensionality of a representational space, given their role in efficient encoding and recognition. By far the most popular tool in such research has been principal component analysis (PCA) because of its relative simplicity and because of its early success as a method for automatic face recognition (Sirovich & Meytlis, 2009; Turk & Pentland, 1991). Briefly, PCA achieves dimensionality reduction by projecting stimuli from their original domain (i.e., pixel-based coordinates) onto a markedly smaller space of uncorrelated dimensions or features—to be clear,

“features” refer here to global whole-face image structures corresponding to the dimensions of face space (e.g., “eigenfaces,” in the case of PCA). Extensive work has shown that PCA can capture some important aspects of psychological face space, as reflected in a range of perceptual tasks, including similarity rating and identification (Dailey & Cottrell, 1999; Hancock, Bruce, & Burton, 1998; Lacroix, Murre, Postma, & van den Herik, 2006; O’Toole, Phillips, Cheng, Ross, & Wild, 2000), as well as expression recognition (Calder, Burton, Miller, Young, & Akamatsu, 2001; Dailey, Cottrell, Padgett, & Adolphs, 2002), race recognition (Furl, Phillips, & O’Toole, 2002), and gender recognition (Graf, Wichmann, Bülhoff, & Schölkopf, 2006).

However, PCA is but one of a larger family of methods that can be used to convert high-dimensional representations into lower-dimensional ones (Zhao, Chellappa, Phillips, & Rosenfeld, 2003). A comparison of alternative

Corresponding Author:

Adrian Nestor, Center for the Neural Basis of Cognition, Carnegie Mellon University, 4400 Fifth Ave., Pittsburgh, PA, 15213-3890
E-mail: anestor@andrew.cmu.edu

architectures based on these methods can be highly informative and desirable for a number of reasons. First, some of these methods promote encoding principles valuable for recognition, such as the statistical independence of the features enforced through independent component analysis (ICA) or the discriminatory power of the features for face identification enforced through linear discriminant analysis (LDA). Second, several projection methods, including ICA (Bartlett, Movellan, & Sejnowski, 2002) and LDA (Belhumeur, Hespanha, & Kriegman, 1997), are able to outperform PCA in automatic face recognition, at least in certain conditions (but see Delac, Grgic, & Grgic, 2005). Last, PCA applications to face perception have certain limitations, such as oversensitivity to low-level pictorial properties (Hancock et al., 1998), and these limitations may be overcome by alternative approaches.

Of particular relevance is the comparison between PCA and ICA (for early work on this comparison in face perception, see Hancock, 2002). In the study of early vision, ICA has become widely successful by accounting for basic processes, such as edge filtering and color opponency (Bell & Sejnowski, 1997; Lee, Wachtler, & Sejnowski, 2002). Assuming that low-level visual representations can be approximated by ICA encoding, it may seem surprising that higher-level visual representations, such as face representations, would revert to a less constraining encoding schema approximated by PCA—feature decorrelation implemented by PCA does not achieve the full statistical independence targeted by ICA (Bartlett, 2007).

In the present work, we investigated the psychological plausibility of three candidate architectures of facial-identity representation based on PCA, LDA, and ICA, respectively. Identifying the method that accounts best for behavioral data can provide key insights into the architecture of face space. To this end, we used well-established behavioral paradigms to collect a comprehensive data set of similarity and identification judgments for pairs of carefully selected face images. Human data were then compared with objective face differences as quantified by PCA, LDA, and ICA. Critically, we implemented a method of feature selection that allowed us (a) to obtain a more principled estimate of the fit of each method with the behavioral data, (b) to compare the relative dimensionality of different face-space architectures, and (c) to rank features based on their explanatory power with regard to human performance.

Last, we considered the importance of color information for face representations by constructing color-based and luminance-based face spaces and assessing the difference between them. Because faces can display complex color patterns that are part of people's common experience with faces as a category, it is important to

determine whether and how color information contributes to representations of facial identity. Thus, we aimed to clarify fundamental properties of face space with regard to its featural makeup, its statistical structure, and its informational content.

Method

Participants

Forty-four Caucasian adults (23 females, 21 males; age range: 19–34 years) with normal or corrected-to-normal vision participated in the experiments in exchange for payment or course credit. Informed consent was obtained prior to participation.

Stimuli

The stimulus set consisted of 480 color images (240 facial identities \times two expressions: neutral and happy) extracted from multiple face databases (Fig. 1). Particular care was taken in the selection and processing of the stimuli to ensure their homogeneity (see Supplemental Method and Results in the Supplemental Material available online). Both the homogeneity and the size of the stimulus set were intended to prevent face comparisons based on gross image dissimilarity or idiosyncratic differences. A quarter of the stimuli were used for behavioral testing, whereas the entire stimulus set was used for constructing candidate face spaces.



Fig. 1. Examples of neutral (top row) and happy (bottom row) face stimuli used in the two experiments. The original, unaltered versions of the images shown here were selected from the Radboud Faces Database (Langner et al., 2010).

Procedure

On each trial, participants were presented with pairs of faces and asked either to rate their similarity on a 5-point scale (Experiment 1) or to decide whether the faces belonged to the same or to a different individual (Experiment 2). An equal number of participants ($n = 22$) took part in each experiment. The two tasks offer complementary advantages in that the former provides direct estimates of similarity, whereas the latter is a simpler, more natural task.

In Experiment 1, stimuli were presented side by side until the participant pressed a key; in Experiment 2, stimuli were presented sequentially, each for 400 ms, and then the participant made a response. In all other respects, the two experiments followed the same design. Each trial started with a 100-ms fixation cross, and each stimulus subtended approximately $3.1^\circ \times 4.6^\circ$. On any given trial, each pair of faces consisted of a neutral version of an individual and the happy expression of the same or another individual. The use of different expressions was intended to preclude equating the task to single-image matching and, thus, to minimize reliance on low-level cues uncharacteristic of everyday recognition. Each participant viewed at least one version of each pair of facial identities and completed 2,000 trials across two 1-hr sessions, each split into five experimental blocks (with short breaks between blocks). Behavioral responses (similarity ratings from Experiment 1 and accuracy from Experiment 2) were encoded into confusability matrices recording the relationship between all identity pairs.

Facial-feature computation

All images were converted to Commission Internationale de l'Éclairage (CIE) 1976 $L^*a^*b^*$ color space, which best approximates the color-opponent properties of human vision (Brainard, 2003). Briefly, the L^* component of the space encodes lightness, whereas the a^* and b^* components encode red:green and yellow:blue ratios, respectively.

Two versions of the stimuli, one using only L^* information and the other using complete $L^*a^*b^*$ color information, were subjected to PCA, LDA, and ICA—see Supplemental Method and Results in the Supplemental Material for more information. The three methods delivered different sets of features (Fig. 2) associated with different candidate face spaces. In each space, a given stimulus can be represented as a vector of coefficients associated with each of these features, and the similarity between stimuli in a pair can be quantified based on these coefficients (here, we quantified similarity using a city-block metric). Such differences were computed across all pairs of stimuli.

Computation of fits to empirical data

To estimate model fits with empirical data, we separately averaged pairwise similarity ratings (Experiment 1) and accuracy-based identification scores (Experiment 2) across participants and correlated these results with model-based face distances. Because only certain subsets of features are likely to reflect perceptual features of human recognition, we implemented a recursive method for feature selection.

First, features were left out one by one, and the outcome of each feature's elimination was assessed by recomputing face distances with the aid of the remaining features and by correlating these distances with the behavioral data. Second, we discarded the feature that produced the best fit among the remaining features, and third, we repeated this process until all features were discarded. Over the course of this process, we initially noticed an increase in the size of the fits (presumably because uninformative features were eliminated first) followed by a decrease (due to the elimination of informative features).

This method yielded a ranking of the explanatory power of the features as well as an approximation of the optimal subset of features (i.e., the one that maximized the correlation with behavioral data). These computations were performed using only half of the experimental stimuli, whereas the second half was used to compute unbiased estimates of fit based on optimal feature subsets. Cross-validation was repeated 20 times using random split halves.

Results

Model fits to behavioral data in Experiment 1 (Fig. 3a) showed main effects of architecture (PCA, LDA, ICA), $F(2, 38) = 160.78, p < .001, \eta^2 = .54$, and color ($L^*, L^*a^*b^*$), $F(1, 19) = 234.85, p < .001, \eta^2 = .20$, as well as an interaction between these two factors, $F(2, 38) = 33.36, p < .001, \eta^2 = .05$. Additional pairwise comparisons showed that ICA outperformed both PCA, $t(19) = 10.75, p < .001$, and LDA, $t(19) = 21.63, p < .001$, whereas PCA outperformed LDA, $t(19) = 6.66, p < .001$. Also, the use of color led to an advantage for all architectures ($ps < .001$), although this advantage was markedly larger for PCA and LDA than for ICA.

Model fits were next compared with estimates of the average consistency of single subjects' similarity ratings and accuracy with the behavioral group data. These estimates were computed by correlating the data of each participant with an average based on the combined data of the other participants. This procedure was repeated for each of the 20 stimulus sets used during cross-validation. The comparison between model and participant

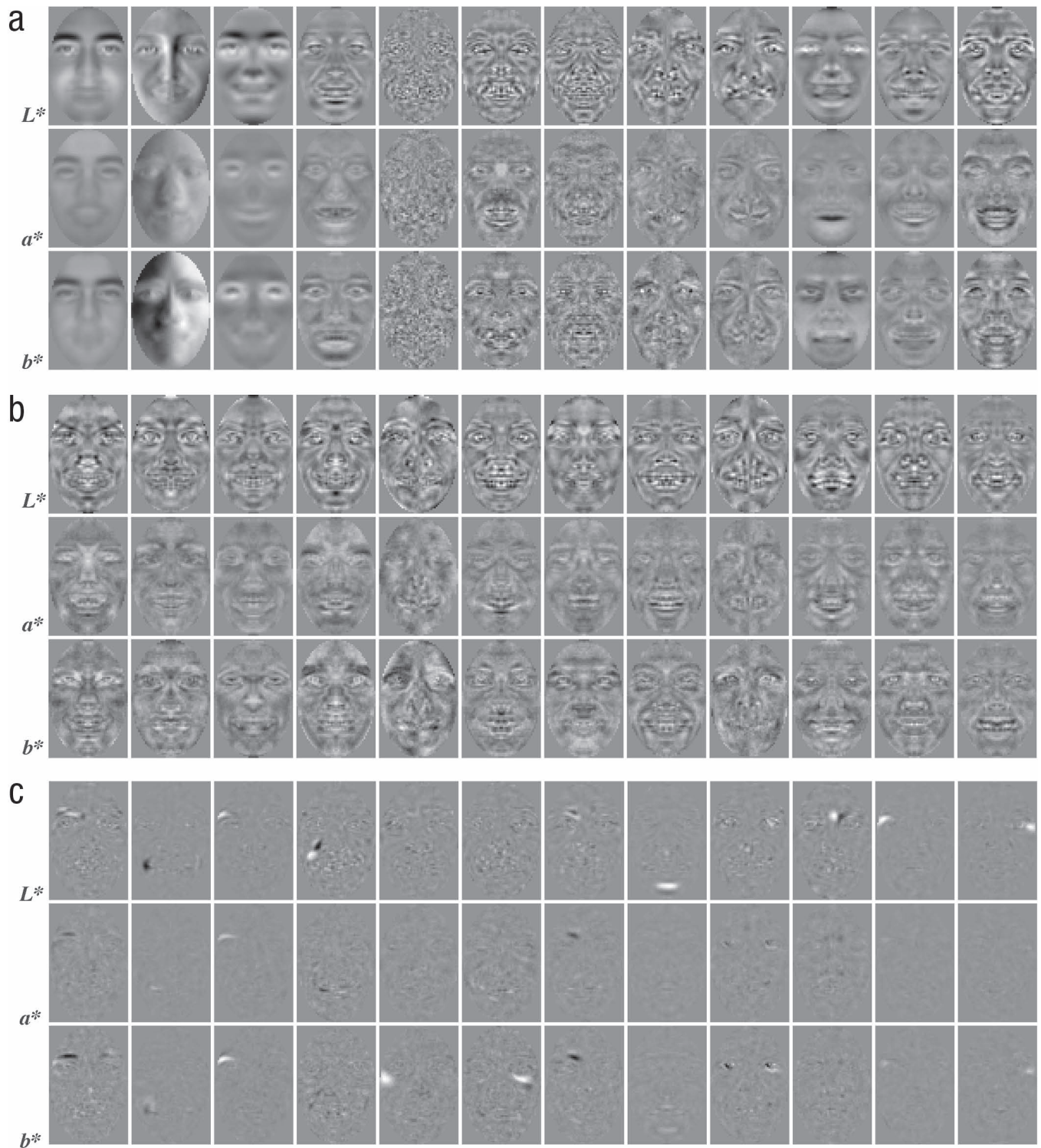


Fig. 2. Examples of color-based face features derived through principal component analysis (i.e., eigenfaces), (b) linear discriminant analysis (i.e., Fisherfaces), and (c) independent component analysis. The three color components, Commission Internationale de l'Éclairage (CIE) 1976 L^* , a^* , and b^* , are separately presented for each feature; the components correspond to the three channels of human vision—lightness, red:green, and yellow:blue. Features are ordered on the basis of their explanatory performance (most to least from left to right) with respect to behavioral data from Experiment 1.

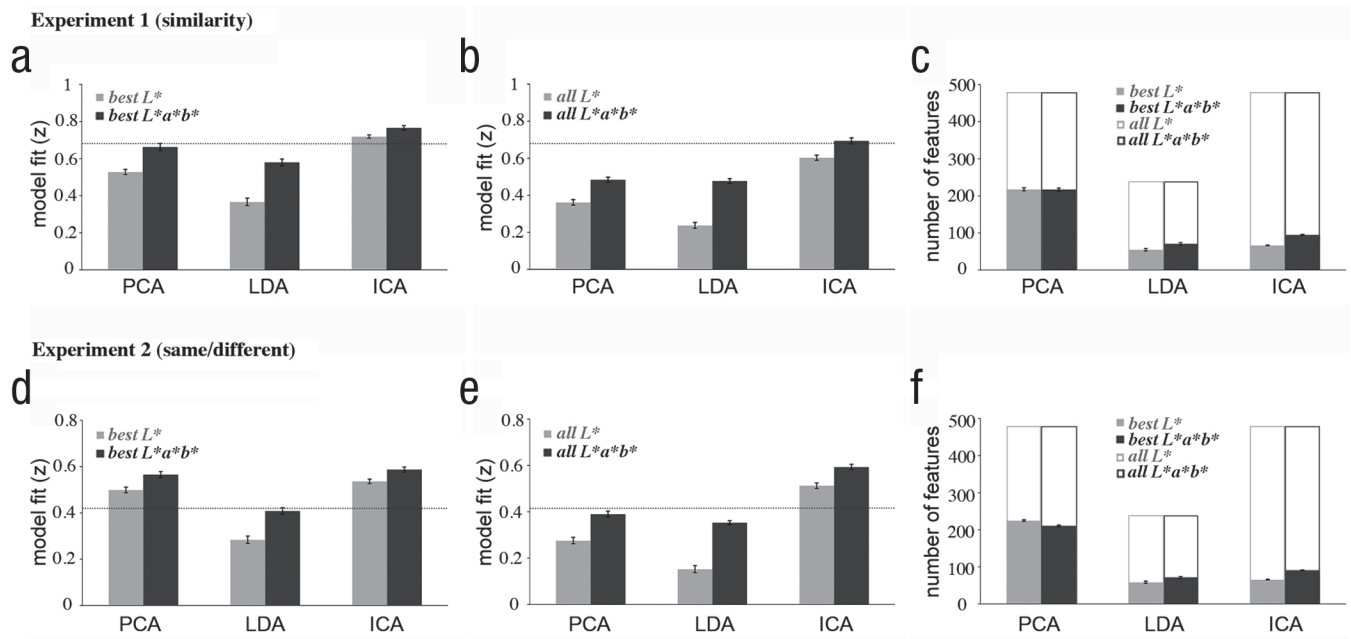


Fig. 3. Results of Experiment 1 (top row) and Experiment 2 (bottom row). The graphs in (a) and (d) show model fits based on optimal sets of features, separately for principal component analysis (PCA), linear discriminant analysis (LDA), and independent component analysis (ICA), both for models including only the L^* component and for models including the L^* , a^* , and b^* components. The graphs in (b) and (e) show fits of PCA, LDA, and ICA models that included all features, both for models including only the L^* component and for models including the L^* , a^* , and b^* components. The graphs in (c) and (f) show the average number of features required by each type of model (for optimal feature subsets) as well as the total (initial) number of features. Dotted lines mark estimates of intersubject consistency. Error bars show ± 1 SE.

consistency showed that ICA but not other architectures surpassed participant consistency, $t(19) = 4.15$, $p < .001$; that is, the ICA model approximated the group data better than the data from actual single subjects did.

The advantage of ICA over other architectures and the advantage of color over lightness alone were replicated in Experiment 2 (Fig. 3d) although, this time, both PCA and ICA architectures surpassed participant consistency ($ps < .001$). The advantages noted for optimal subsets of features were also replicated using entire sets of features, further attesting to the robustness of our results. However, in these cases, model fits were markedly smaller (Figs. 3b and 3e).

Next, a comparison of space dimensionality showed that both ICA and LDA required smaller feature sets than PCA (Figs. 3c and 3f). The reduction in dimensionality relative to the original size of the feature set was particularly notable for ICA (about six times smaller). Finally, an examination of ICA features showed that, as expected, they were sparser than PCA or LDA features (Fig. 2). That is, few pixels had large positive or negative values, whereas the remaining ones were close to zero as a result of ICA's reliance on a sparse source model (Bartlett et al., 2002). However, we note that ICA features were not limited to local information—for instance, many encoded symmetrical facial structures corresponding to the eyes or

the nostrils. Thus, despite their sparsity, ICA features can, in principle, encode configural information.

For additional analyses of the behavioral data, aimed at quantifying their robustness and consistency, see Supplemental Method and Results, Figure S1, and Figure S2 in the Supplemental Material.

Discussion

What type of perceptual code underlies human face representations, and what general principles govern its organization? To address these issues, we compared the ability of three face-space architectures based on PCA, LDA, and ICA, respectively, to account for behavioral data. These architectures involve considerable differences in the statistical properties of face-space structure and in the featural makeup of internal representations. Our results point to the superior explanatory power of ICA-based features by showing that an ICA architecture accounts for the properties of human face space to a larger extent and with a more parsimonious feature base than its alternatives. These results are significant in two main respects.

First, they provide evidence for statistically independent features in face encoding and, implicitly, for the role of information maximization as implemented by ICA (Bartlett, 2007)—here, the maximization of the amount of

information that representations carry about objective stimulus properties. Previously, these ideas have been explored theoretically in relation to perceptual phenomena such as the other-race effect and face adaptation (Tanaka, Kantner, & Bartlett, 2012; Webster & MacLeod, 2011). Our results yielded empirical support for the validity of this theoretical framework in the study of face perception—for instance, if statistical independence can account for fine-grained individual recognition, then its extension to more general effects (e.g., the other-race effect) becomes all the more plausible.

Second, the present findings may help to bridge the study of high-level visual recognition and that of low-level perception. Because early visual processes such as edge filtering and color opponency (Bell & Sejnowski, 1997; Lee et al., 2002) are relatively well explained by ICA, our results lend credence to the hypothesis that statistical independence describes a general property of visual representations rather than a domain-specific or stage-specific one.

A further issue examined here concerns the role of color in face perception. Whereas representations of facial identity are largely thought to discard color information (Kemp, Pike, White, & Musselman, 1996; Yip & Sinha, 2002), recent results point to the use of color in face detection (Bindemann & Burton, 2009) and gender categorization (Nestor & Tarr, 2008). The present findings suggest that the use of color extends to facial identity, as we noted a systematic color advantage for all architectures and experimental manipulations.

Face encoding and recognition are highly flexible processes (Griffin, McOwan, & Johnston, 2011; Mielle, Caldara, & Schyns, 2011) that require a robust representational base. The present results suggest that color-based independent features can serve this function. As an important caveat, though, we note that our data were acquired in tightly controlled experimental settings both with respect to the nature of the tasks and of the stimuli tested. Thus, future research is necessary to establish the generality of our findings by exploring a variety of tasks as well as more naturalistic stimuli exhibiting other types of variability (e.g., due to lighting or viewpoint).

In conclusion, our work yields support for an ICA-based account of face space and for the role of color in the representation of facial identity. These results prompt a reconsideration of fundamental aspects of the structure and informational content of face space. More generally, they provide new evidence for efficient information encoding grounded in the structure of the visual input in the domain of face perception.

Acknowledgments

We thank Gary Cottrell and the other members of the Perceptual Expertise Network for helpful comments.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding

This research was supported by the National Science Foundation (Grant BCS0923763 to M. B. and D. C. P. and Grant SBE-0542013 to M. B. from the Temporal Dynamics of Learning Center).

Supplemental Material

Additional supporting information may be found at <http://pss.sagepub.com/content/by/supplemental-data>

References

- Bartlett, M. S. (2007). Information maximization in face processing. *Neurocomputing*, *70*, 2204–2217.
- Bartlett, M. S., Movellan, J. R., & Sejnowski, T. J. (2002). Face recognition by independent component analysis. *IEEE Transactions on Neural Networks*, *13*, 1450–1464.
- Belhumeur, P. N., Hespanha, J. P., & Kriegman, D. J. (1997). Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *19*, 711–720.
- Bell, A. J., & Sejnowski, T. J. (1997). The “independent components” of natural scenes are edge filters. *Vision Research*, *37*, 3327–3338.
- Bindemann, M., & Burton, A. M. (2009). The role of color in human face detection. *Cognitive Science*, *33*, 1144–1156.
- Brainard, D. H. (2003). Color appearance and color difference specification. In S. K. Shevell (Ed.), *The science of color* (pp. 191–216). Washington, DC: Optical Society of America.
- Calder, A. J., Burton, A. M., Miller, P., Young, A. W., & Akamatsu, S. (2001). A principal component analysis of facial expressions. *Vision Research*, *41*, 1179–1208.
- Dailey, M. N., & Cottrell, G. W. (1999). Organization of face and object recognition in modular neural network models. *Neural Networks*, *12*, 1053–1074.
- Dailey, M. N., Cottrell, G. W., Padgett, C., & Adolphs, R. (2002). EMPATH: A neural network that categorizes facial expressions. *Journal of Cognitive Neuroscience*, *14*, 1158–1173.
- Delac, K., Grgic, M., & Grgic, S. (2005). Independent comparative study of PCA, ICA, and LDA on the FERET data set. *International Journal of Imaging Systems and Technology*, *15*, 252–260.
- Furl, N., Phillips, P. J., & O’Toole, A. J. (2002). Face recognition algorithms and the other-race effect: Computational mechanisms for a developmental contact hypothesis. *Cognitive Science*, *26*, 797–815.
- Graf, A. B., Wichmann, F. A., Bülthoff, H. H., & Schölkopf, B. (2006). Classification of faces in man and machine. *Neural Computation*, *18*, 143–165.
- Griffin, H. J., McOwan, P. W., & Johnston, A. (2011). Relative faces: Encoding of family resemblance relative to gender means in face space. *Journal of Vision*, *11*(12), Article 8. Retrieved from <http://www.journalofvision.org/content/11/12/8.abstract>

- Hancock, P. J. B. (2002). Psychological correlates of face shapes. In J. A. Bullinaria & W. Lowe (Eds.), *Proceedings of the Seventh Neural Computation and Psychology Workshop: Connectionist models of cognition and perception* (pp. 145–155). Singapore: World Scientific.
- Hancock, P. J. B., Bruce, V., & Burton, M. A. (1998). A comparison of two computer-based face identification systems with human perceptions of faces. *Vision Research*, *38*, 2277–2288.
- Kemp, R., Pike, G., White, P., & Musselman, A. (1996). Perception and recognition of normal and negative faces: The role of shape from shading and pigmentation cues. *Perception*, *25*, 37–52.
- Lacroix, J. P., Murre, J. M., Postma, E. O., & van den Herik, H. J. (2006). Modeling recognition memory using the similarity structure of natural input. *Cognitive Science*, *30*, 121–145.
- Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud faces database. *Cognition & Emotion*, *24*, 1377–1388.
- Lee, T. W., Wachtler, T., & Sejnowski, T. J. (2002). Color opponency is an efficient representation of spectral properties in natural scenes. *Vision Research*, *42*, 2095–2103.
- Miellat, S., Caldara, R., & Schyns, P. G. (2011). Local Jekyll and global Hyde: The dual identity of face identification. *Psychological Science*, *22*, 1518–1526.
- Nestor, A., & Tarr, M. J. (2008). Gender recognition of human faces using color. *Psychological Science*, *19*, 1242–1246.
- O'Toole, A. J. (2011). Cognitive and computational approaches to face recognition. In A. J. Calder, G. Rhodes, M. H. Johnson, & J. V. Haxby (Eds.), *The Oxford handbook of face perception* (pp. 15–30). New York, NY: Oxford University Press.
- O'Toole, A. J., Phillips, P. J., Cheng, Y., Ross, B., & Wild, H. A. (2000). Face recognition algorithms as models of human face processing. In *Proceedings of the 4th IEEE International Conference on Automatic Face and Gesture Recognition* (pp. 552–557). Los Alamitos, CA: IEEE Computer Society.
- Sirovich, L., & Meytlis, M. (2009). Symmetry, probability, and recognition in face space. *Proceedings of the National Academy of Sciences, USA*, *106*, 6895–6899.
- Tanaka, J. W., Kantner, J., & Bartlett, M. (2012). How category structure influences the perception of object similarity: The atypicality bias. *Frontiers in Psychology*, *3*, 147. Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3368386/>
- Turk, M., & Pentland, A. (1991). Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, *3*, 71–86.
- Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, *43*, 161–204.
- Webster, M. A., & MacLeod, D. I. (2011). Visual adaptation and face perception. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*, 1702–1725.
- Yip, A. W., & Sinha, P. (2002). Contribution of color to face recognition. *Perception*, *31*, 995–1003.
- Zhao, W., Chellappa, R., Phillips, P. J., & Rosenfeld, A. (2003). Face recognition: A literature survey. *ACM Computing Surveys*, *35*, 399–458.