Selective Attention Among Presumed Classifiers in the Human Face Recognition System

Peter Kalocsai & Irving Biederman

University of Southern California
Hedco Neuroscience Building
Los Angeles, CA 90089-2520

kalocsai@selforg.usc.edu
ib@rana.usc.edu

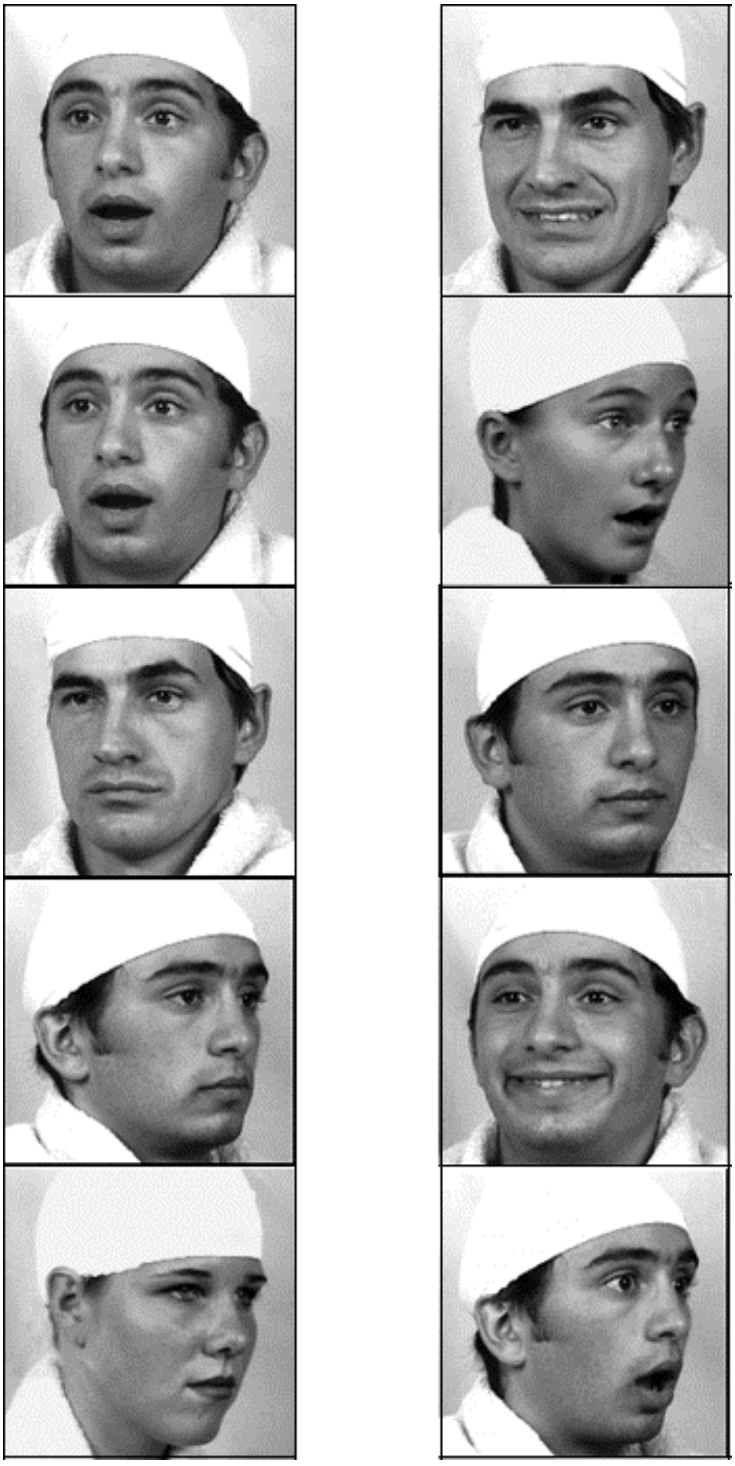
Poster presented at the Meeting of the Association for Research in Vision and Othalmology, Fort Lauderdale, FL., May, 1995.

The authors thank Professor Christoph von der Malsburg for his help in making the Face Recognition System available, to Jean-Marc Fellous for his management of the FRS, and to József Fiser for his help in processing the face images. This research was supported by the McDonnell-Pew Foundation Program in Cognitive Neuroscience grant T89-01245-029 and the U. S. Army Night Vision and Electronics Sensors Directorate (NVESD), ARO grant No. DAAH04-94-G-0065.

Overview

Purpose. Theorists have posited that separate modules code for different aspects of faces. To what extent can people selectively attend to one of these variables while ignoring the others? **Method.** Subjects judged whether a pair of brief (100 msec), masked, sequential presentation of faces were of the same individual, expression (neutral, smiling, surprised), gender, or orientation (0°-60°). The hair and clothing were occluded so the subjects had to rely on the face itself. (Which faces are of the same individual? Which are female? Answers below.) The image similarity of each pair of faces was assessed by a model (Buhman, Lades, and von der Malsburg, 1990) with units approximating the activation of a lattice of V1 simple cells. This allowed specification of physical (image) similarity across the qualitatively differing variables. On this scale, for example, the image variation for orientation was large relative to that produced by gender. **Results.** Variation in an irrelevant attribute (or attributes) generally interfered with the speed and accuracy in judging two images to be the same, indicating a failure of selective attention. Gender was the exception to this pattern. Overall, performance correlated highly with the similarity values calculated by the model, so that greater similarity produced shorter RTs and lower error rates. That is, the magnitude of an effect of an irrelevant variable on a judgment could generally be predicted from the magnitude of its image variation. Gender judgments were an exception to this pattern. **Conclusion.** A model of V1 simple cell similarity space predicts judgments of identity, orientation and expression, indicating that the presumed modules, if they do exist for these attributes, cannot be accessed independent of the original V1 image variation.

Which of the following pairs of faces (left and right) are the same or different individuals?



Now go back and judge which pairs have the same expression? Have the same orientation? Have the same gender?

Purpose

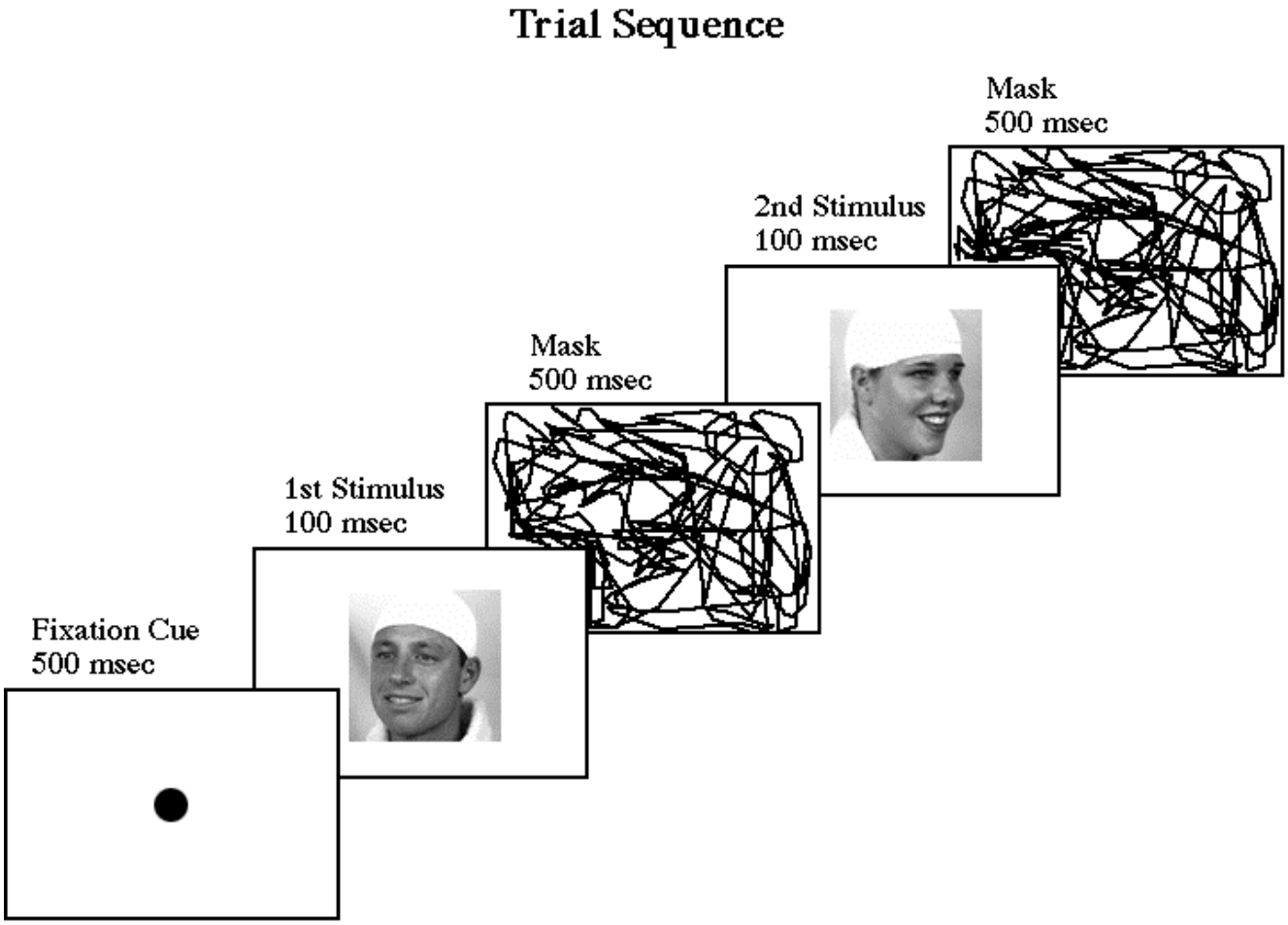
There is some evidence for separate neural loci for processing of facial expression and identity (Hasselmo et al, 1989; Perrett et al., 1984), orientation (Perrett et al., 1992) and gender (Bruce, et al., 1987).

- To what extent can facial characteristics - Identity, Expression, Orientation, and Gender - be judged independent of variation of the other attributes?

- As irrelevant attributes differ the face images become less similar. To what extent is performance predictable from a simple cell similarity space (discussed later)?

General Method

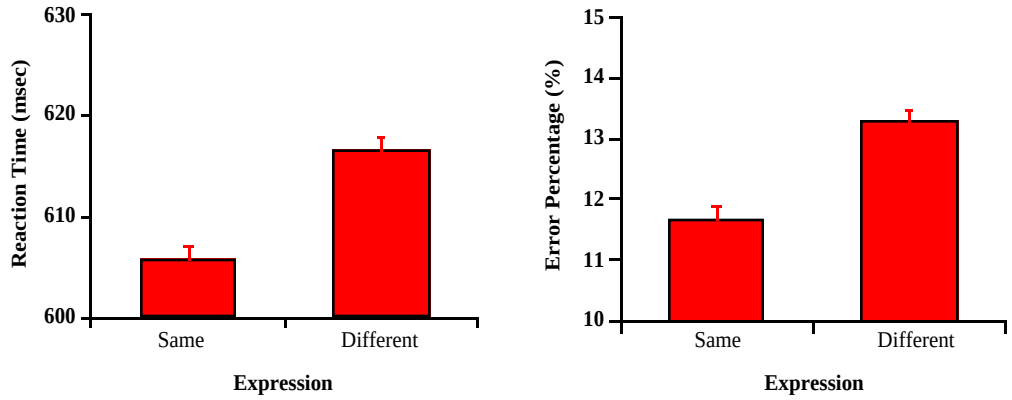
In separate experiments subjects judged whether two faces (same age, no hair or clothing or easy features) were of the same or different individual, had the same or different gender, expression, or orientation. The faces were viewed briefly (100 msec) and sequentially (with masks after each image). The two faces could differ in orientation by 0 to 60 deg (from 20 deg left to 40 deg right) and emotional expression (neutral, happy, surprise).



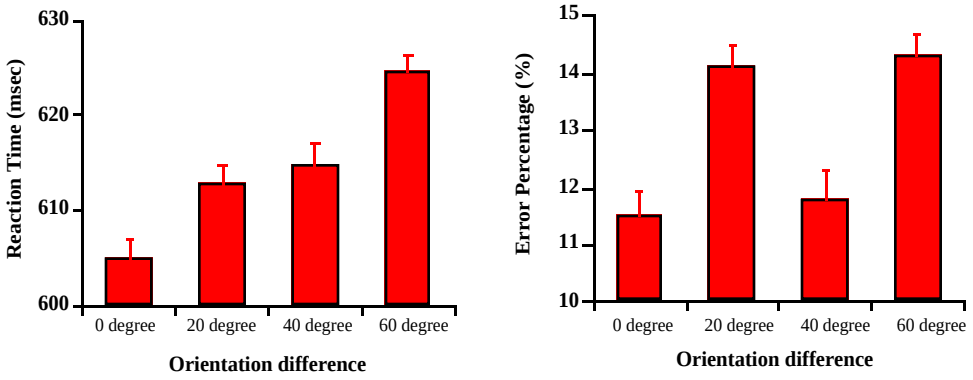
Judging Identity

Variation in Expression, Orientation, and Gender all interfered with judgment of Identity as shown below.

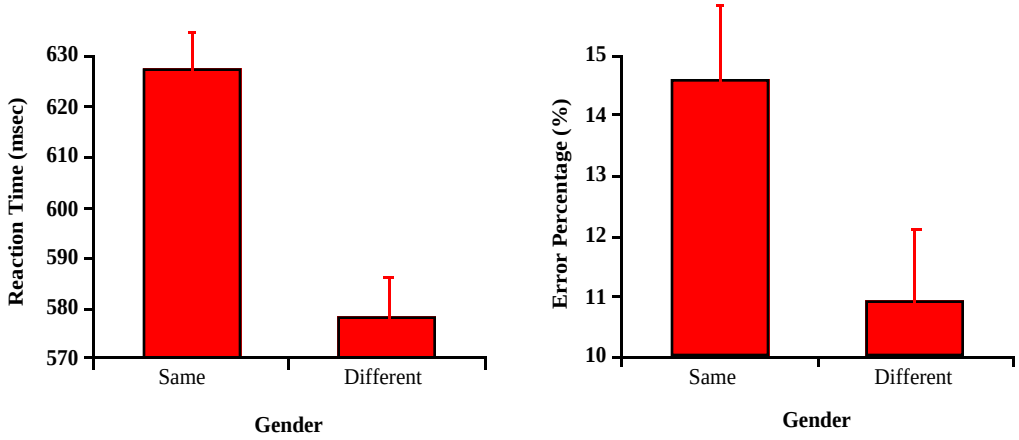
Effect of Expression on Identity



Effect of Orientation on Identity



Effect of Gender on Identity*

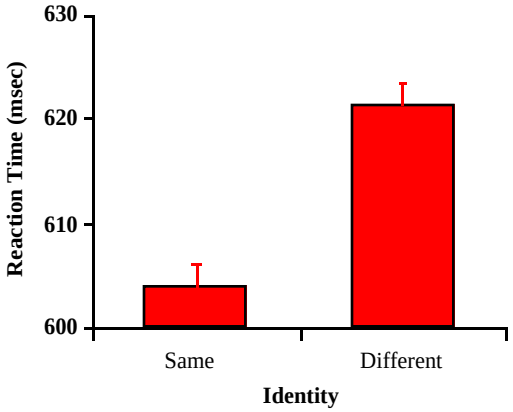


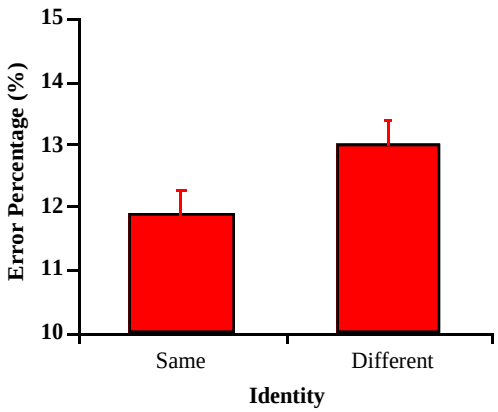
* Presented data is based on different identity trials.

Judging Expression

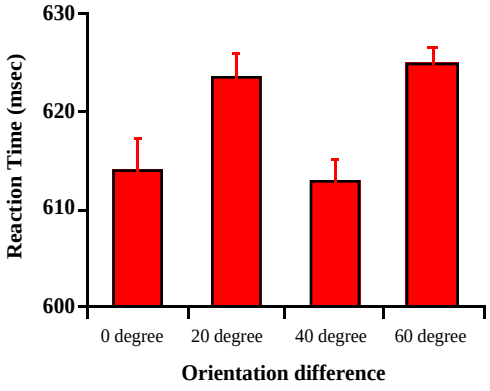
Variation in Identity, Orientation, and Gender all interfered with judgment of Expression as shown below.

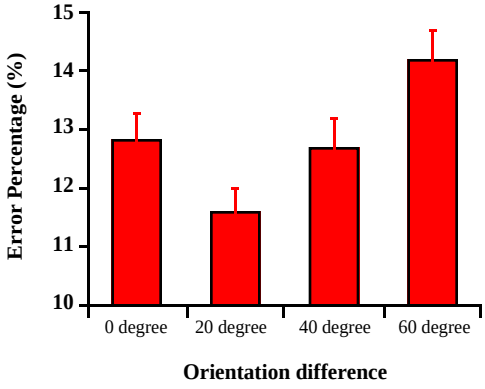
Effect of Identity on Expression



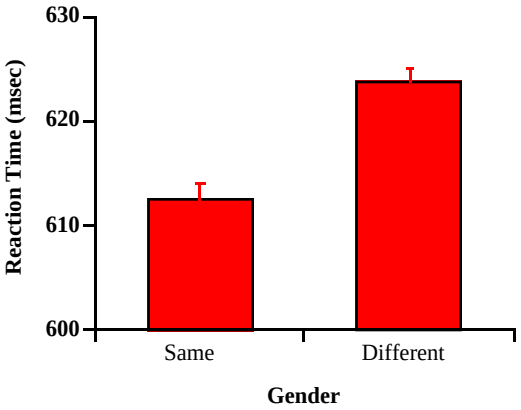


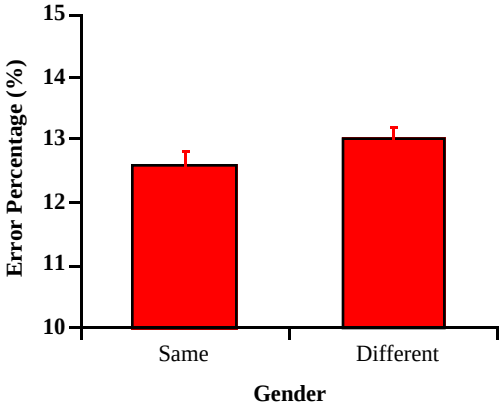
Effect of Orientation on Expression





Effect of Gender on Expression

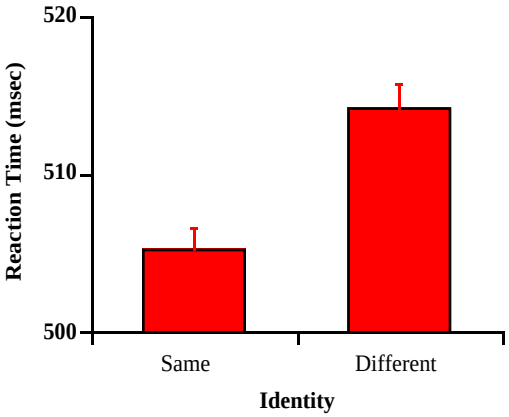


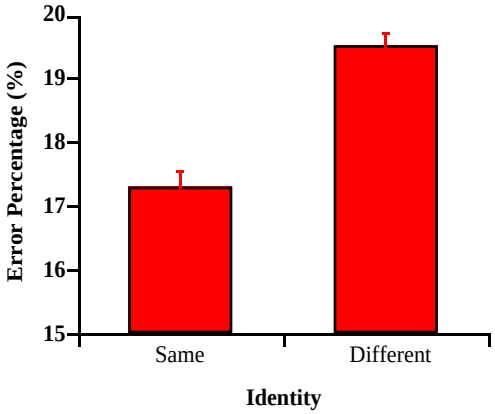


Judging Orientation

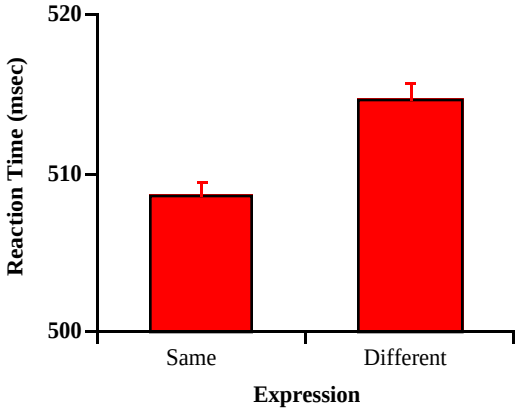
Variation in Identity, Expression, and Gender all interfered with judgment of Orientation as shown below.

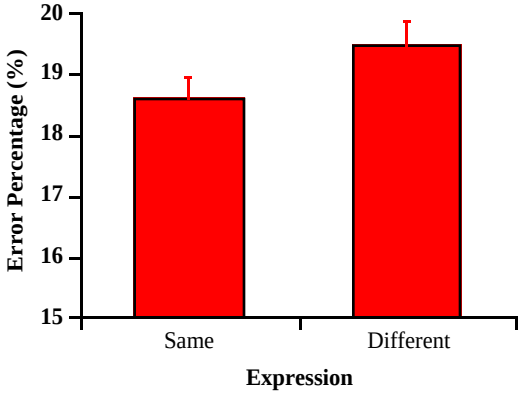
Effect of Identity on Orientation



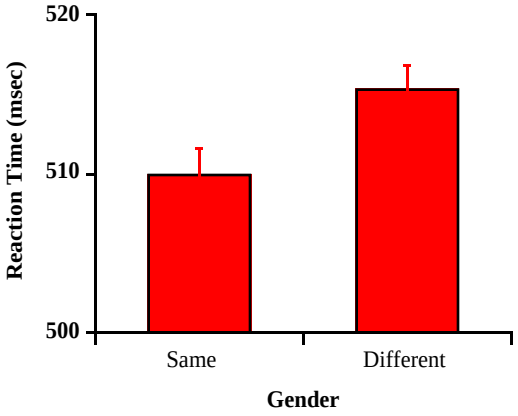


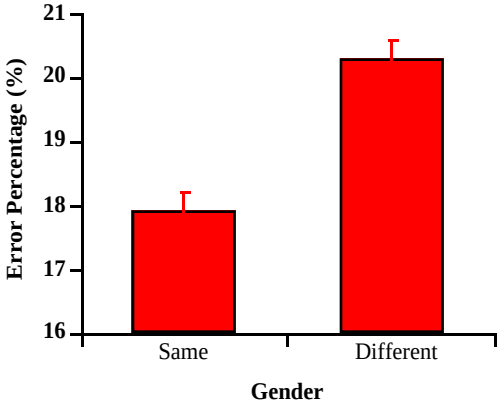
Effect of Expression on Orientation





Effect of Gender on Orientation

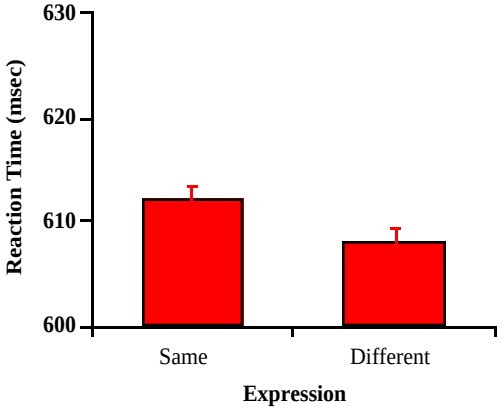


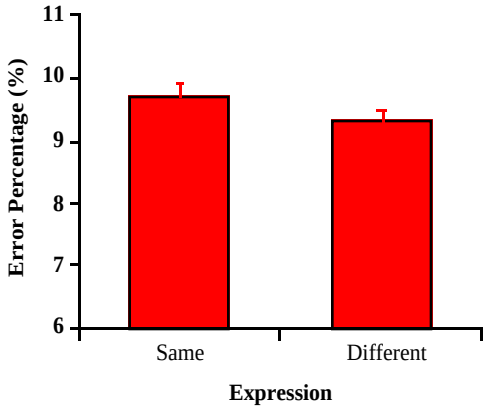


Judging Gender

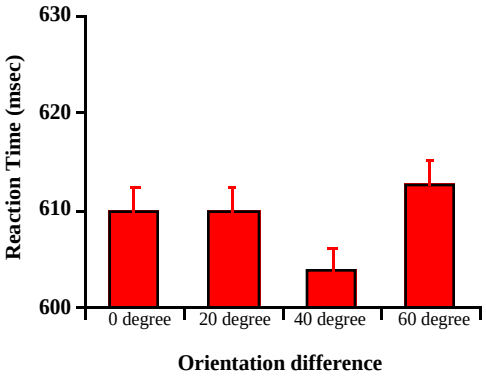
Variation in Identity, Expression, and Orientation had no effect on judgment of Gender. This result is remarkable because image variation produced by gender differences were small relative to that produced by the other variables.

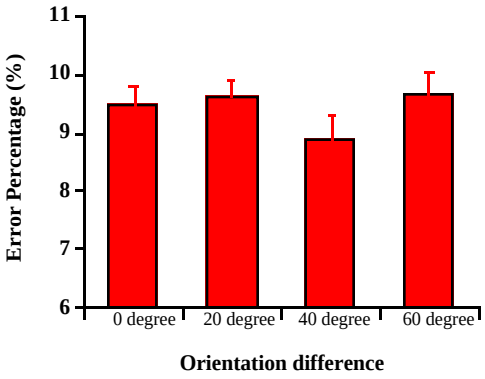
Effect of Expression on Gender



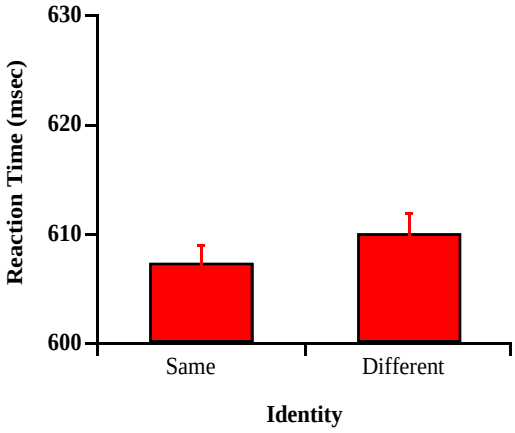


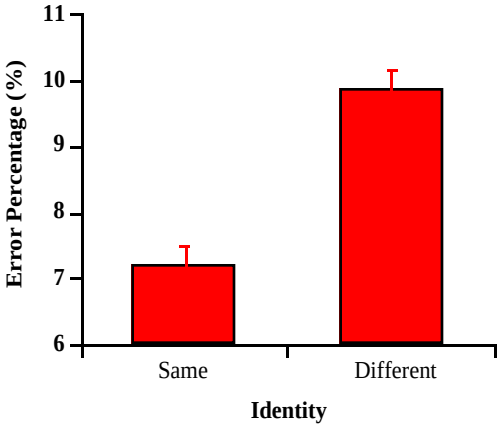
Effect of Orientation on Gender





Effect of Identity on Gender





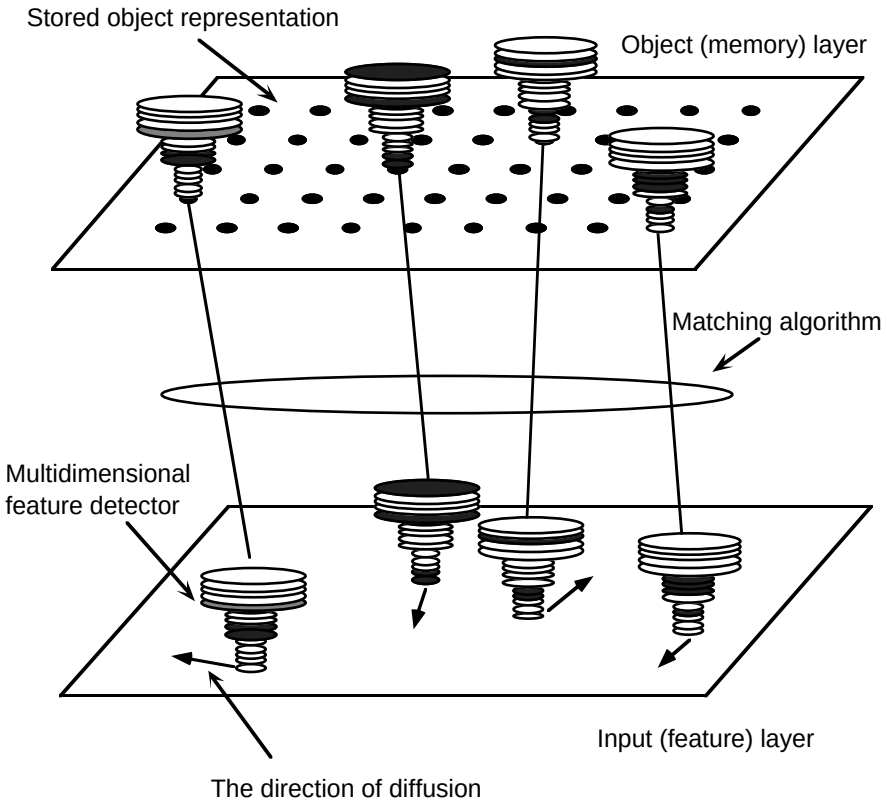
Simple Cell Similarity Space of Images of faces

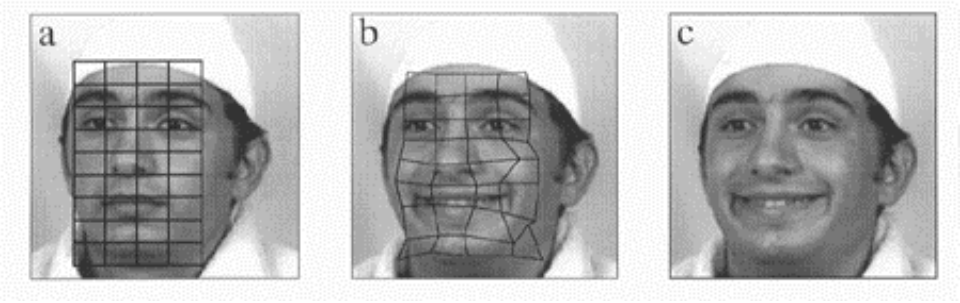
The simple cell similarity of the images was determined with the Buhmann, Lange, & von der Malsburg's (1990) simple cell (SC) model for face recognition.

General Structure of the SC Model

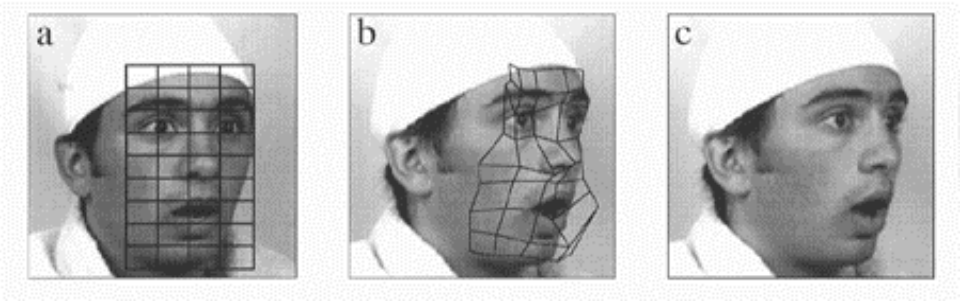
Architecture

The SC model first convolves each input image with a set of Gabor kernels at five scales and eight orientations arranged in a 5 x 9 lattice (figure below). The positioning of the lattice over an image is shown in the left hand column, labeled (a) of the figure with faces below. The set of kernels at each node in the lattice is termed a “Gabor jet.” The activation values of the kernels in each jet along with their positions are stored for each of the images to form a “gallery”.

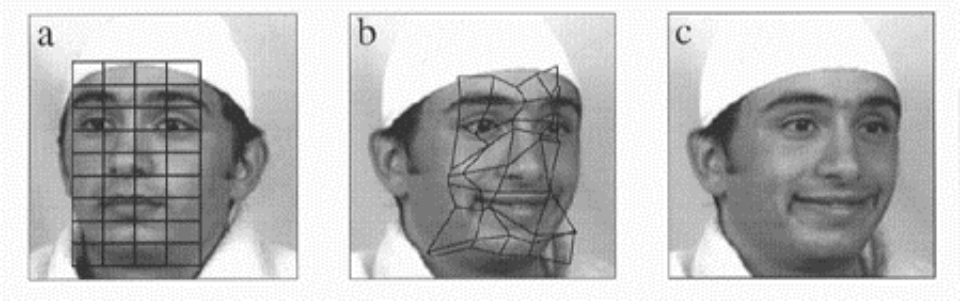




Same orientation-different expression



Different orientation-same expression



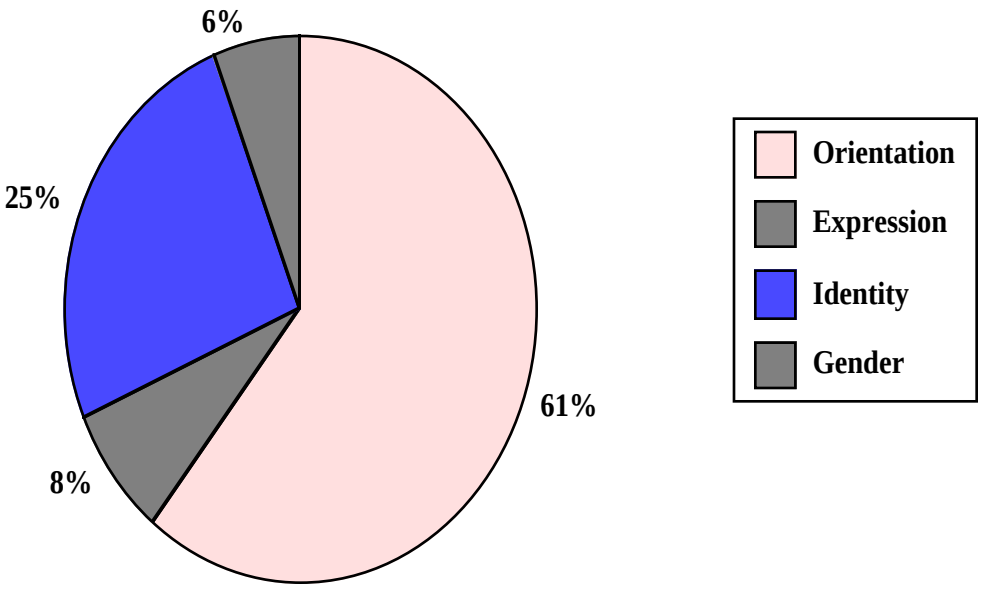
Different orientation-different expression

Matching (Determination of Similarity)

The similarity between a test image and the various stored images (gallery) is calculated by stochastic optimization that allows each of the jets to diffuse (gradually change its position) to optimize the similarity in kernel values and distances relative to adjacent jets. The result of the diffusion over a pair of faces is shown in the middle column (b) in the above figure (the test faces without the distorted grids are presented in the right hand column (c)). To the extent that the jets move independently, the resultant positions will no longer produce a rectangular lattice, as illustrated in the figure. In general the more distorted the lattice, the less the similarity of the image to the original. The most similar match of the test image is interpreted to be the recognition response of the model. The model achieves 83% accuracy in correctly recognizing a second image of an individual (out of a gallery of 160 individuals), even with considerable variation in facial expression but only slight differences in orientation. When the correct face does not receive the highest rank, it is almost always among the next two faces.

Distribution of variance in the stimuli

A calculation of the contribution of each variable to the total image variation of the stimulus set showed that gender difference produced the smallest magnitude of image variation and orientation the greatest.

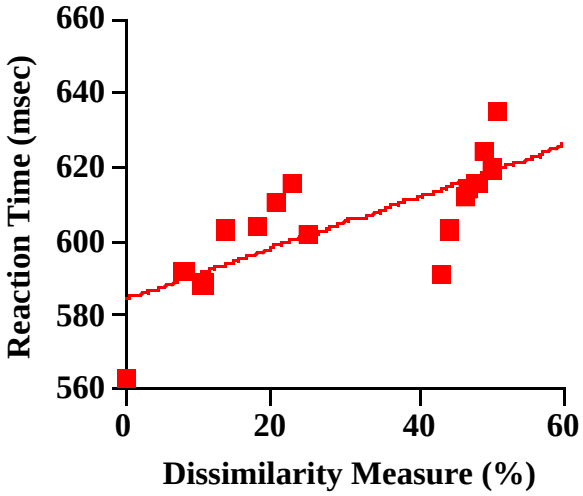


Regression Functions

The regression functions for identity, expression, and orientation have very similar positive slopes indicating that image similarity had approximately the same effect on these variables. Gender was an exception to this general rule in that image dissimilarity had no effect on gender judgments.

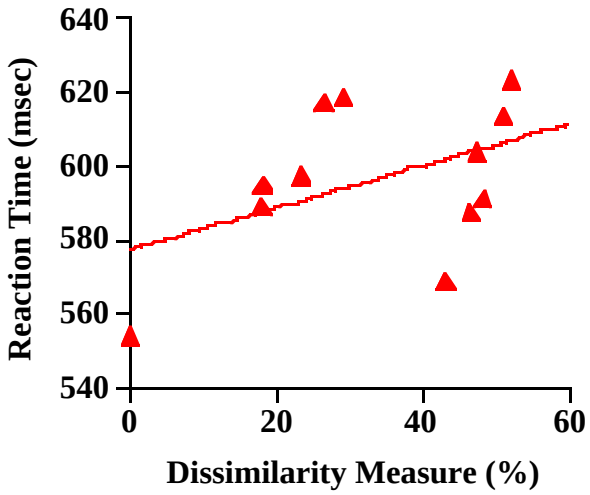
Identity

r = .73



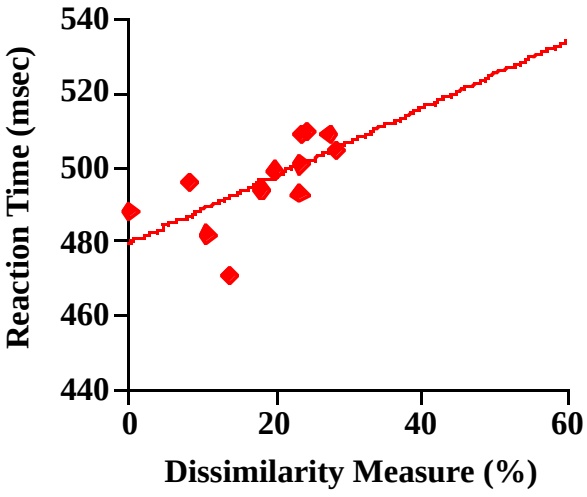
Expression

r = .54



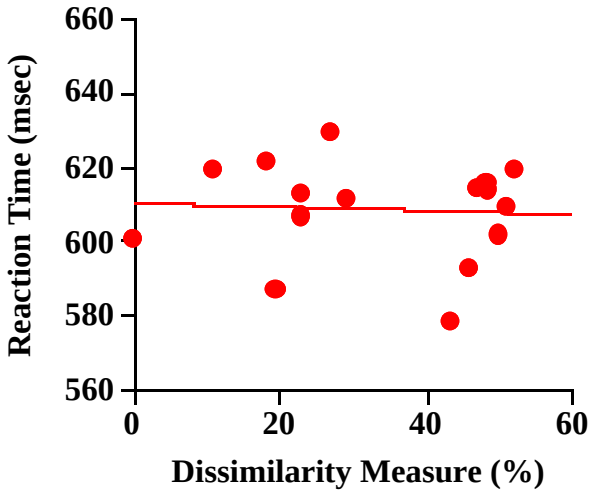
Orientation

r = .67



Gender

r = -.05



Conclusions

- Selective attention is only possible for gender; judgment of identity, expression and orientation are all affected by variation in the other attributes.

- A model of V1 simple cell similarity space predicts judgment performance of identity, orientation and expression, but not gender.

References:

Bruce, V., Ellis, H., Gibling, F. & Young, A. (1987). Parallel processing of the sex and familiarity of faces. *Canadian Journal of Psychology, 41,* 510-520.

Buhmann, J., Lades, M., & von der Malsburg, C. (1990). Size and Distortion Invariant Object Recognition by Hierarchical Graph Matching. In *International Conference on Neural Networks,* (pp. 411-416). San Diego.

Hasselmo, M. E., Rolls, E. T., & Baylis, G. C. (1989). The role of expression and identity in the face-selective responses of neurons in the temporal visual cortex of the monkey. *Behavioural Brain Research, 32,* 203-218.

Lades, M., Vortbrüggen, J. C., Buhmann, J., Lange, J., von der Malsburg, C., Würtz, R. P., & Konen, W. (1993). Distortion Invariant Object Recognition in the Dynamic Link Architecture. *IEEE Transactions on Computers, 42,* 300-311.

Perrett, D. I., Smith, P. A. J., Potter, D. D., Mistlin, A. J., Head, A. S., Milner, A. D., & Jeeves, M. A. (1984). Neurones responsive to faces in the temporal cortex: Studies of functional organization, sensitivity to identity, and relation to perception. *Human Neurobiology, 3,* 197-208.

Perrett, D. I., Hietanen, J. K., Oram, M. W., & Benson, P. J. (1992). Organization and functions of cells responsive to faces in the temporal cortex. *Philosophical Transactions of the Royal Society of London B,* 335, 23-30.