

PSYCHOPHYSICS OF FACE RECOGNITION

Stephen Link
University of California, San Diego, USA
McMaster University, Hamilton, Canada
link@ucsd.edu

Kang Lee
University of Toronto
Toronto, Canada
kang.lee@utoronto.ca

Abstract

Same-different judgments show that in face recognition experiments mean response times and correct response proportions are similar to previous psychophysical studies of sameness and difference. These new experimental results and analyses show that the “Interval of Uncertainty” remains fixed regardless of the overall accuracy of performance. Furthermore, improvements in performance occur without providing feedback about correct or error responses, suggesting that the image of a face continually sharpens as a result of repeated presentations.

“And mark the fleers, the gibes, and notable scorns that dwell in every region of his face” Iago implores the Moor Othello. As seen in this use of facial expressions by Shakespeare, the face acts as a projector of inner emotional states, providing a surface of interest to playwrights, painters, lyricists, poets, criminologists, and even psychologists. Faces convey emotions such as agitation, grief, loss, excitement, and joy with such clarity that the very naming of these emotions conjures up images of faces bearing corresponding expressions. No doubt the study of such facial images reaches back into prehistory, but our interest is in the current excitement about face recognition and analysis, often described as facial cognition (Wenger and Townsend, 2001).

Perhaps the introduction of face analysis, of biometrics, by Sir Francis Galton may be taken as a starting point for the modern study of faces. Galton (1878) combined in a single image the photographs of eight faces. The result was an average “face exhibiting the features most common to the collection of (in this case criminal) faces. Galton commented that “I begin by collecting photographs of the persons with whom I propose to deal. ... Then by a simple contrivance I make two pin-holes in each of them, to enable me to hang them up one in front of the other, like a pack of cards, upon the same pair of pins, in such a way that the eyes of all the portraits shall be as nearly as possible superimposed. ... A composite portrait represents the picture that would rise before the mind’s eye of a man who had the gift of pictorial imagination in an exalted degree.”

Judgments about faces with respect to criminality (in Galton’s case), family similarity, beauty, gender, or age, are common in daily life and in experiments designed to investigate facial cognition. But face recognition requires judgments about faces that are primarily judgments of sameness or difference. Such judgments occur as routinely as judgments of largeness or smallness but are based on a different principle, such as matching to a sample or an ideal. Such judgments of similarity require a very different treatment, a different theoretical basis, for their analysis than do judgments of largeness or smallness.

Quite a number of years intervened between the development of Fechner’s original model for the comparative judgment of stimuli and the recognition that some of the judgments made in psychological experiments were not judgments of differences in stimulus magnitude at all but were, instead, judgments of equality or sameness. Urban (1907) proposed that three different responses may be employed in the typical discrimination experiment using a fixed standard and variable comparison stimuli. A comparison stimulus may be judged to be

DISCRIMINATING VISUAL ANGLES
CARTWRIGHT 1941

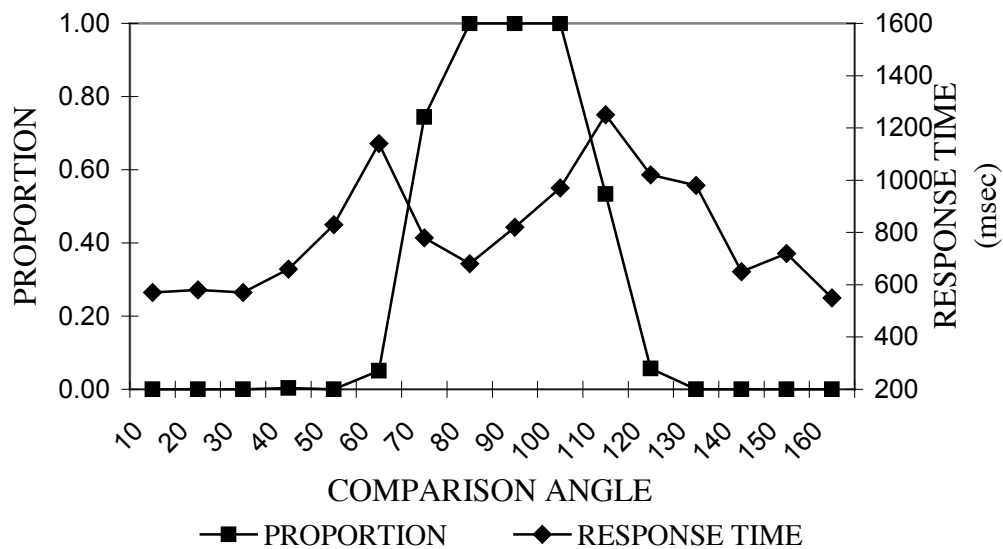


Figure 1: Results from Cartwright (1941) on judging whether a test angle was within the interval of 60° to 100°. Mean response times in milliseconds.

smaller, larger, or equal to a standard stimulus. Although experimental subjects were able to use such response categories when making judgments between two stimuli, the theory of how such judgments might occur required many years of subsequent development (cf. Link, 1992). In short, judgments of smaller or larger are one type of judgment and judgments of equality or inequality, of sameness or difference, are the only other type of judgment to be discovered since Fechner’s amazing introduction of comparative judgment in 1860.

The experimental result that differentiates judgments of similarity, or equality, from judgments of largeness or smallness is illustrated in Figure 1. Here are results from a typical method of constant stimuli experiment employing visual angles as stimuli (Cartwright, 1941). Experimental subjects decided if a presented angle varying to 160° from 10° was within a learned interval of 60° to 100°. Both response latency, the time from the presentation of the stimulus to the response, also called response time (RT), and response proportions show the existence of what Urban (1910) defined as an “Interval of Uncertainty.” This “interval” is the contiguous interval within which the probability of making the judgment equal is greater than 0.50. Cartwright’s results show how, at the edges of this “interval,” at the point of 50% responding, the mean response latency, or response time, reaches a peak. Indeed, mean response times appear to be a function of the distance of a comparison stimulus from the edges of the interval of uncertainty. This result is one of many showing that the application of a judgment model based on monotonic changes in performance as a function of stimulus difference is not applicable to judgments of sameness and difference.

Figure 2 presents results from an experiment on discriminating a facial feature. The experimenter (Lee) required 10 subjects to reply SAME or DIFFERENT when a test face was compared against a visible standard face. Only a single dimension of difference, the distance between the eyes, changed by varying the distance in pixels between the eyes (approximately 0.04cm/pixel). Comparison faces changed from a very large 10 pixel distance greater than the Standard to 10 pixels less than the Standard. As is clearly evident, the same pattern of mean

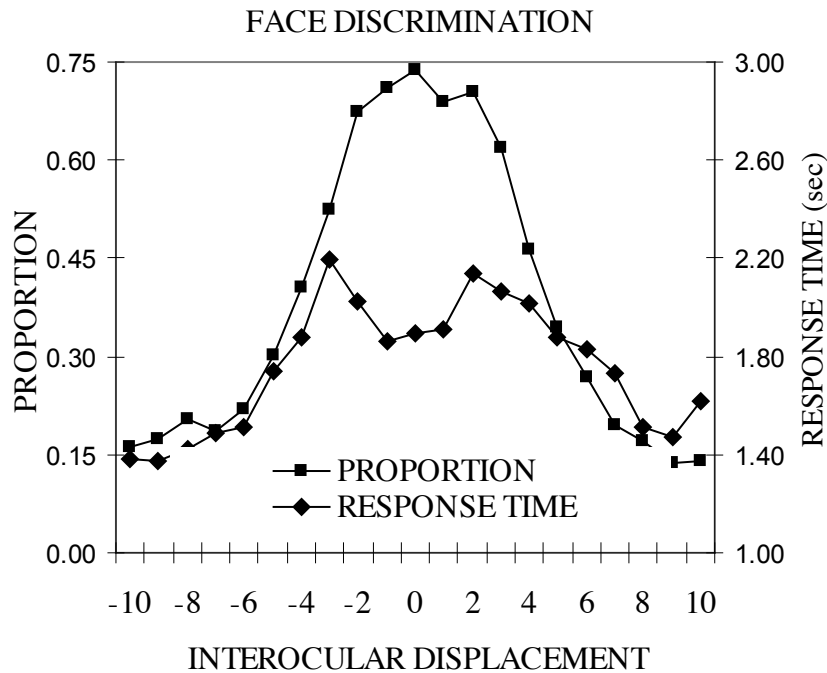


Figure 2. Results for judging the similarity of two simultaneously presented faces. Positive (negative) numbers index increasing (decreasing) number of pixels between eyes. Mean response times in seconds.

response time and response proportions occurs in these face recognitions as in Cartwright's, and many other experiments, when judgments are of sameness or difference.

In the other conditions of this experiment subjects were split into three groups distinguished by the number of seconds of time studying the Standard face before 1000 comparison stimuli were presented. Three groups received 10, 20 or 30 seconds of time studying the Standard face only once at the beginning of the experiment. Comparison stimuli ranged from -10 pixels interocular distance less than the Standard to 10 pixels greater than the Standard, as in the case of simultaneous presentations of the Standard and Comparison face. The Standard occurred as a comparison on a random 50% of trials, each of the remaining comparisons presented on a random 2.5% of trials. Thus there were 500 presentations of the studied face, the Standard, and 500 presentations of faces differing in the number of pixels between the eyes. Following the study period the experiment began by presenting a face on a computer controlled display in front of the seated subject. The subject responded by using keys on the left and right-hand sides of a computer keyboard to indicate SAME or DIFFERENT. Response choice and the time between the onset of the face display and the choice, that is, the response latency or response time, were measured. There was no feedback.

Response times and response choices (SAME or DIFFERENT) were analyzed to determine whether subjects followed instructions. Two subjects failed to meet a criterion of mean response times less than 3 seconds and were removed from these analyses. One discovery was that responses prior to 500 milliseconds were at chance performance. These trials were removed. Otherwise there are a total of 42,328 trials entering into these results.

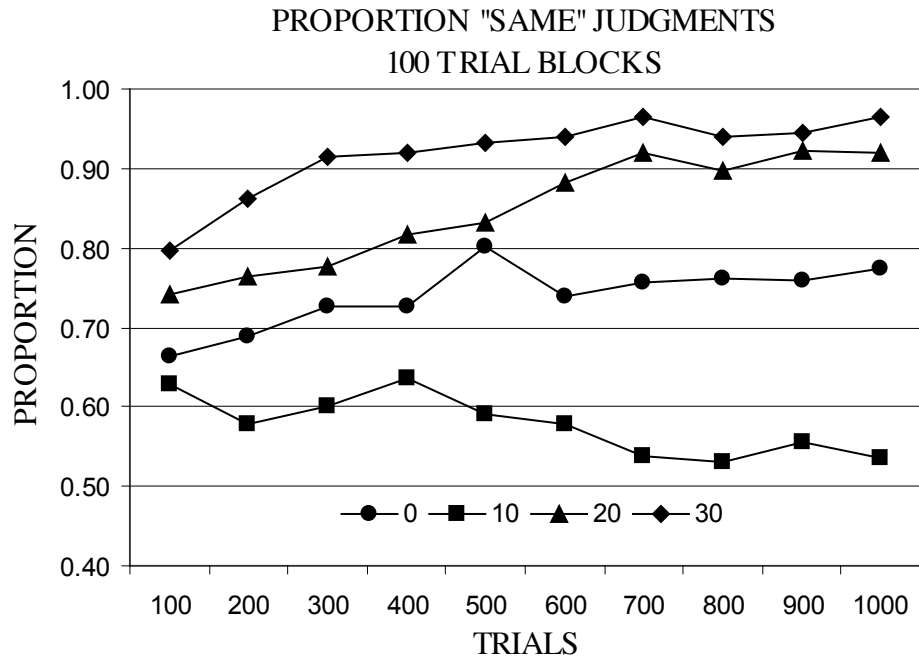


Figure 3. Performance by groups within 100 trial blocks. The Standard occurred at random on 50% of trials. Group 0 (N=10) had simultaneous presentations of Standard and Comparison. Groups 10 (N=11), 20 (N=12), and 30 (N=11) received face study times of 10, 20, and 30 seconds only once, at the beginning of the experiment.

Figure 3 shows the proportion of correct responses when the Standard was presented as a Comparison. In the 20 and 30 second study conditions the subjects differ by initial study time, have no feedback, and yet still improve their identification of the Standard across trials. Performance seems to stabilize from trials 600 to 1000. The 10 second study condition shows a diminution in performance nearly reaching a level of chance performance. Interestingly the simultaneous presentation condition does not lead to better performance than the 20 and 30 second study conditions

An important test concerning the “Interval of Uncertainty” can be constructed from these data. Obviously the various conditions produce differing numbers of correct responses with the conditions ordered from 30 to 20 to simultaneous to 10 seconds study time. Do these variations in performance derive from changes in the width of the “Interval of Uncertainty” or does the “Interval” remain fixed but performance changes due to other factors? By using data from the last 400 trials, where performance seems to have stabilized, Intervals of Equality can be determined and compared.

Results from this analysis appear in Figure 4. The three psychometric functions show the proportion of “SAME” judgments for those conditions where performance improved over trials and then stabilized during the last 400 trials. Given the number of subjects in these groups the functions contain at least 4000 trials each, and may be considered a fair representation of the performance of each group.

The three psychometric functions appear to overlay one another, especially in the range of performance from 0.25 to 0.75. The Interval of Uncertainty appears to span a fixed distance from about -3.5 pixels to +4.0 pixels. Although there is a slight shift to the left of zero in these functions, even taking into account some slight shift, the size of the interval of uncertainty seems fixed. Thus the changes in marginal performance during the last 400 trials are not due to less discriminability as measured by the size of the Interval of Uncertainty.

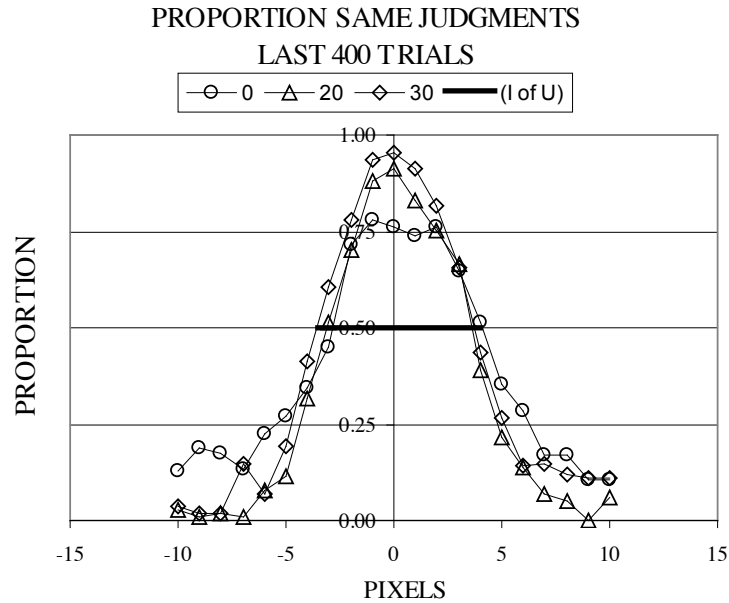


Figure 4. Psychometric “Same” functions for Groups 0, 20 and 30 obtained from trials 601 to 1000. The Interval of Uncertainty extends from -4pixels to +3.5 pixels.

Notice that the heights of these functions at zero are a measure of the marginal proportion of correct SAME responses for the last 400 trials shown in Figure 3.

To confirm the relation between response proportions and response times the same 400 trials were analyzed with respect to response time. In Figure 4 appear the individual Chronometric Functions for conditions 0, 20 and 30 and the average of these mean response times for each pixel distance. Quite obviously the same pattern of increases up to the lower edge of the Interval of Uncertainty, the decrease in response time within the IU, the increase at the right hand edge of the IU and subsequent decrease in mean RT beyond the right hand IU edge are apparent.

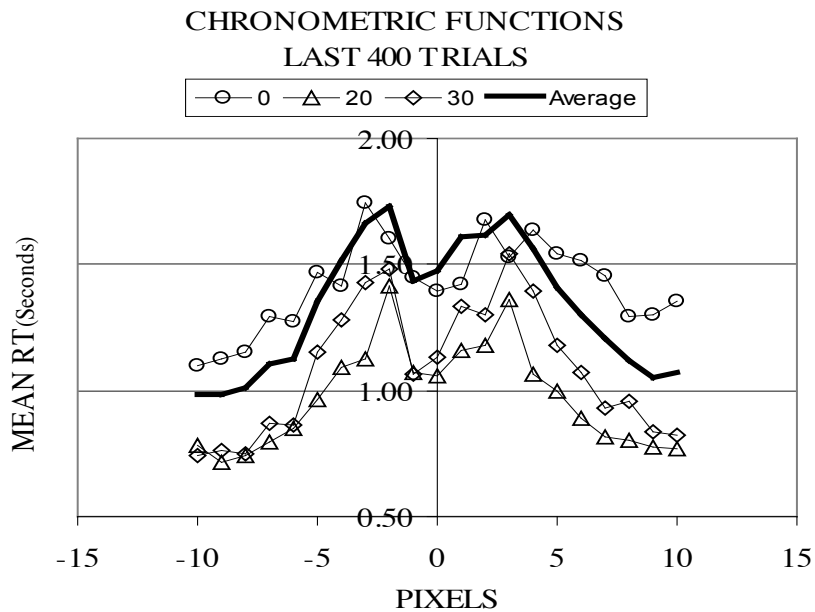


Figure 5. Mean response times (correct and errors together) for the last 400 trials where performance seems to stabilize.

Why are these observations important? First, Urban proposed that the Interval of Uncertainty was a better measure of discrimination than a 75% threshold of responding. A small interval suggested a very fine ability while a large interval showed little discriminability at all. Associated with this idea was the idea that the IU was under voluntary control, could broaden to create poorer performance, or narrow to produce a finer discrimination of Sameness. In this regard the IU was taken to be too arbitrary and not only a measure of discriminability. In these data we see that IU may remain constant as performance changes. This form of constancy, if confirmed in other experiments employing Same-Different judgments, places a strong constraint on a theory of Same-Different judgments.

Second, these observations demonstrate quite clearly that the results of these face recognition experiments correspond to results from other, classic, experiments investigating sameness and difference. Changes in response time vary as a function of distance from the edges of the Interval of Uncertainty not as a function of gross stimulus difference. Thus, the relation between response time and response proportions is not the same relation as that that exists in experiments where judgments of larger or smaller are appropriate. Analytical methods for studying same-different judgments are not those derived from statistical models based simply on decreases or increases in stimulus difference.

Third, the data show quite clearly that performance improves without reinforcement. Repeated presentations of the stimuli seem to create the “composite portrait” suggested by Galton (1878). Yet the diminution in performance in the condition allowing only 10 seconds of study of the Standard face suggests there must be sufficient initial development to create a basis from which the composite may solidify.

Last, the apparent constancy of the Interval of Uncertainty raises the question of what determines the Interval of Uncertainty.

References

- Cartwright, D. (1941). Relation of decision time to categories of response. *American Journal of Psychology*. 54, 174-196.
- Galton, F. (1878). Composite portraits. *Nature*. 97-100. May, 1878.
- Link, S. W. (1992). *The wave theory of difference and similarity*. Lawrence Erlbaum Associates: Hillsdale, New Jersey.
- Urban, F. M. (1907). On the method of just perceptible differences. *Psychological Review*. 14, 244-253.
- Urban, F. M. (1910). On the method of constant stimuli and its generalizations. *Psychological Review*. 17, 229-259.
- Wenger, M. and Townsend, J. T. (2001). *Computational, geometric, and process perspectives on facial cognition*. Lawrence Erlbaum Associates. Mahwah, New Jersey.