Bulletin of the Psychonomic Society
1980, Vol. 15 (4), 269-270

Judgments of facial attractiveness in the absence of eye movements

ALVIN G. GOLDSTEIN and JANET PAPAGEORGE
University of Missouri, Columbia, Missouri 65211

Forty portraits were selected from a larger number of portraits previously rated by five judges on a 6-step scale of attractiveness. In a subsequent study, subjects also rated the 40 faces on the same scale, but viewing time was restricted to either 150 or 1,000 msec per portrait. Even at exposures of 150 msec, correlation of subjects' ratings with judges' ratings was highly reliable.

The results of several laboratory investigations suggest that faces that are considered by observers to be atypical or uncommon along some psychological dimension are more likely to be remembered than faces considered to be ordinary or typical. For example, beauty (Cross, Cross, & Daly, 1971), attractiveness (Fleishman, Buckley, Klosinsky, Smith, & Tuck, 1976; Shepherd & Ellis, 1973), and distinctiveness (Cohen & Carr, 1975) have been shown to be related to face recognition performance. Although the relationship between these findings and the von Restorff effect has been mentioned (Going & Read, 1974), the connection has not been clearly enunciated, perhaps because the relationships between von Restorff's observations and the recognition data are far from clear. Going and Read (1974) referred to Loftus' (1972) work and wondered whether frequency of eye fixation could be used to explain the relationship between recognition memory performance and psychological atypicality. This approach makes good sense because eye movements have rarely been controlled in the investigations of attractiveness and recognition. Typically, investigators have allowed subjects to look at target faces at least 1 sec, and more often up to 5 sec, during which 4 or as many as 20 eye fixations, respectively, could have occurred. Presumably, if an atypical face is more "interesting" than a common face, eye fixation frequency in looking at the former would be greater than that involved in looking at the latter. If Loftus (1972) is correct, this improved recognition could be explained because more information about the atypical face would be encoded as a function of the more frequent eye fixations.

One way to study this question is to reduce the number of eye fixations made while looking at faces that vary in judged typicality. This method would work only if valid attractiveness judgments could be made when observation time was reduced to 150 msec. The method used in the present investigation reduced eye fixation frequency to one fixation per stimulus face by reducing exposure time to 150 msec. (Execution of voluntary eye movements while scanning a display has been shown to take about 250 msec/saccade, including periods of fixation, so that about four eye movements are possible in a 1-sec observation period.) Three separate studies were performed, all addressed to this question: Can subjects make valid (i.e., accurate) judgments of facial attractiveness while viewing a face for only 150 msec?

METHOD

Only one of the three studies will be reported in detail, but the differences between the studies were minor. Portraits used in the three studies were all obtained from the same source. Individual photographs in two of the three studies were identical. Two experimenters performed the three studies.

Subjects

Five undergraduates served as judges in preliminary phases of the research. Sixty men and women undergraduates served as subjects in the experiment. None of the subjects had been judges.

Apparatus

A two-channel mirror tachistoscope (Scientific Prototype) was used in the conventional manner to present 40 black-and-white portraits (original yearbook photographs) of college men and women. Portraits were selected to minimize differences in pose, expression, and so on. Each face subtended a visual angle of approximately 2.5 deg across its greatest dimension. This visual extent was increased for half the subjects by the introduction of a rigidly mounted low-power (2.5X) optical system. Operation of the tachistoscope on each trial was initiated by the experimenter. The fixation stimulus, which preceded each trial, spatially coincided (except for minor variations of a few minutes of arc) with the bridge of the nose on each portrait.

Procedure

More than 800 photographs were sorted by five judges, working independently, into six categories of attractiveness (1 = exceptionally handsome or beautiful; 2 = very handsome or beautiful; 3 = handsome or beautiful; 4 = plain but not homely; 5 = homely; 6 = ugly). From this large pool of rated portraits (obtained range = 1.2-6.0), pictures with greatest interjudge agreement were selected to form two closely matched sets of 20 photos each. Pictures within these sets differed by .2-step interval on the attractiveness scale. Instructions informed the subjects to look into the tachistoscope and to report to the experimenter their readiness to view a portrait. The experimenter then pressed the start switch, which first activated the fixation field for 500 msec, then illuminated a portrait for

0090-5054/80/040269-02$00.45/0

either 150 or 1,000 msec, depending on the experimental condition, and, finally, returned the visual field to its original unilluminated condition. A visual mask was not used. Half the subjects were tested on one set of photographs, and the remaining subjects judged the second set of portraits. One randomly determined presentation order for each set of photographs was maintained for all subjects.

Four groups of subjects viewed "normal-sized" (approximately 2.5 deg) or "large" (approximately 5 deg) portraits for either 150 or 1,000 msec. After looking at each portrait, subjects rated the face along the attractiveness dimension, using the rating scale described above. Rating instructions to subjects and judges were identical.

RESULTS AND DISCUSSION

Data were analyzed by treating the portraits as the unit of analysis. Table 1 shows for each of the four visual size and exposure time combinations the correlations between the judges' mean ratings and the subjects' mean ratings. All correlations were significant ($p < .01$). Clearly, valid attractiveness ratings can be made when subjects have only 150 msec to observe a face.[1] This means eye movements are not necessary for perceiving beauty in a face. This generalization holds even when the image extends beyond the area of clear vision (large, 150 msec, $r = .70$). Differences among the sizes of the correlations are instructive. As expected, longer looking time results in greater agreement between the judges' and the subjects' ratings, but because the instructions to subjects in the 1,000-msec conditions were ambiguous, the larger correlations might be due to eye movements and not to simple processing time. Unfortunately, we have no measure of the test-retest reliability of the attractiveness ratings, so we do not know if judge-subject correlations are asymptotic for 1,000-msec exposure. Without this information, we cannot decide whether 1,000 msec is the minimum time needed to make valid (i.e., equal to judges with unlimited time) attractiveness ratings.[2]

Perhaps a few additional words should be said about the similarity of outcomes for the two 150-msec conditions. Although some portion of the retinal images in both the large and small retinal size conditions fell outside the area of clearest vision, more of the large image had to be unclear to the subjects. Yet, judgments of attractiveness remained unaffected by this change in size of the retinal image. Many reasons for this fact come to mind, but the most likely explanation is the following. With 150-msec exposure time, eye movements were impossible, but without a visual mask following the face stimulus, processing of the visual input continued (probably for approximately 250 msec, but this estimate is debatable) after the physical stimulus disappeared. Even though some part of the larger stimulus was unclear, enough information was perceived that subjects could accurately rate the face. After all, the ratings were not precise in the first place, and although the correlations were high, they were not perfect.[3]

Table 1
Correlation (r) of Judges' and Subjects' Attractiveness Ratings

Inspection Time (in Milliseconds)				
Small Retinal Image		Large Retinal Image		
150	1,000	1,000	150	Total r
.68	.83	.77	.70	.79

REFERENCES

COHEN, M. E., & CARR, W. J. Facial recognition and the von Restorff effect. *Bulletin of the Psychonomic Society,* 1975, **6**, 383-384.

CROSS, J. F., CROSS, J., & DALY, J. Sex, race, age, and beauty as factors in recognition of faces. *Perception & Psychophysics,* 1971, **10**, 393-396.

FLEISHMAN, J. J., BUCKLEY, M. L., KLOSINSKY, M. J., SMITH, N., & TUCK, B. Judged attractiveness in recognition memory of women's faces. *Perceptual and Motor Skills,* 1976, **43**, 709-710.

GOING, M., & READ, J. D. Effects of uniqueness, sex of subject, and sex of photograph on facial recognition. *Perceptual and Motor Skills,* 1974, **39**, 109-110.

LOFTUS, G. Eye fixations and recognition memory for pictures. *Cognitive Psychology,* 1972, **3**, 525-551.

SHEPHERD, J. W., & ELLIS, H. D. The effect of attractiveness on recognition memory for faces. *American Journal of Psychology,* 1973, **86**, 627-633.

SIMPSON, W. E., & CAPETANOPOULOS, C. Reliability of smile judgments. *Psychonomic Science,* 1968, **12**, 57.

SIMPSON, W. E., & CRANDALL, S. J. The perception of smiles. *Psychonomic Science,* 1972, **29**, 197-200.

NOTES

1. Related findings have been reported by Simpson and Crandall (1972). Judgments of the magnitude of smiles are almost unchanged by increasing exposure time from 20 to 120 msec. Recognition of the faces displaying the smiles, however, was only slightly better than chance at the 20-msec exposure, and it was well above chance levels at 120 msec.

2. Reliability magnitude estimates of smile judgments were found to be quite robust over a 1-month period (Simpson & Capetanopoulos, 1968), suggesting that, perhaps, judgments of attractiveness will be found to be highly reliable.

3. Immediately following the rating task, subjects in all three studies were given an unannounced recognition memory test, the results of which have not been included in this report. Omission was prompted by these facts: In two of these studies recognition memory performance was totally uncorrelated with attractiveness ratings when exposure was limited to 150 msec. When exposure was increased to 1,000 msec in one experiment, attractiveness was still unrelated to recognition performance, a finding that was not in keeping with the results reported by other investigators. Additional problems arose when the results of a third study clearly indicated that recognition performance was related to attractiveness ratings with observation time set at 150 msec. Throughout all of this disquieting variability in the results of the three studies, two findings never changed: Attractiveness ratings made by judges with unlimited time correlated significantly with ratings made by subjects who viewed the faces for 150 msec, and recognition performance remained at a remarkably high level even though the study time was only 150 msec.

(Received for publication February 8, 1980.)