The perception of smiles*

W. E. SIMPSON and SALLY J. CRANDALL†
State University College at Oswego, Oswego, N.Y. 13126

Os made magnitude estimates of the smiles on photographs of faces exposed for various durations. Very brief exposures (.02 sec) resulted in discriminability of smiles equivalent to that obtained with unlimited exposure. Repeated presentation of a face did not alter the magnitude estimates, and the high level of discriminability under brief exposures occurred even when recognition of the faces (old-new) was poor. Perception of a smile appears to occur extremely rapidly, perhaps as a wholistic process.

There is currently a surge of interest in the human face as a stimulus for psychological processes. Research with infants, for example, suggests that the face is a stimulus with unique potential for controlling visual attention and early social processes (Lewis, Kagan, & Kalafat, 1966; Haaf & Bell, 1967). Memory for faces is known to be extraordinarily good (Yin, 1969; Smith & Nielsen, 1970; Galper & Hochberg, 1971). And there is a long and sustained interest in faces as indicators of emotion (Woodworth & Schlosberg, 1954; Ekman, 1972). All of these studies presuppose questions about the *perception* of faces, a topic about which we know very little.

How are facial expressions "read" by an observer? Two possible types of processing are template matching and feature analysis. In the former type, facial expressions are read as *gestalts* or wholes. For example, the components of a smile (upturned mouth, visible teeth, wrinkles around the eyes?) are processed simultaneously by comparing the face against a template of features labeled "smile." Support for such a process can be obtained by showing that facial expressions are perceived in an all-or-none fashion, i.e., once a critical duration of exposure to the face has been reached, sufficient for perception, further increments in exposure duration will have no effect upon the judgment of the expression. For a short-term memory task, Bradshaw & Wallace (1971) predicted that drawings of faces presented in the "normal mode" (upright photographic positives) would be processed in this way, as *gestalts*. Smith & Nielsen (1970) had Os compare schematic faces in a same-different judgment task and found evidence to support the template matching hypothesis as long

as interstimulus intervals were very short (1 sec).

The alternative position, feature analysis, suggests that facial expressions are perceived by a time-dependent process. The short-term memory literature provides considerable evidence in favor of this position. For example, Smith and Nielsen's study showed that with interstimulus intervals on the order of 10 sec, the time required for a "same" judgment, comparing one stimulus face with a face held in memory, increased with the number of features that were identical for the two faces. The longer interstimulus interval condition would seem more characteristic of usual situations of facial perception, and therefore the implication is that faces are perceived by sequential comparison of features. In contrast to their prediction, Bradshaw and Wallace's data also supported the feature analysis position. Their O searched an array of faces to find a match for a memorized standard. The fewer the distinguishing features of the target face, the longer was the search time. Bradshaw and Wallace, therefore, concluded that the comparison of faces was done by a process of serial feature analysis, although they also stressed the rapidity with which this was carried out. Each face required only 1 to 2 sec of processing time.

There are two possible problems in extrapolating the above conclusions about memory to making predictions about perception. The studies following the memory paradigm have asked O to indicate whether or not specific faces presented at two points in time were identical. In the perception paradigm of interest here, the O is asked to make a graded judgment about facial expressions. Standards for judging facial expressions are probably complex and somewhat unique to each perceiver. Thus there is no reason to expect that the perception of a facial expression necessarily involves the same kind of processing as does identification of a face.

The second possible problem concerns the stimuli in use. The memory studies cited above have involved schematic or drawn faces in which a set of fixed features is combined in various ways. If, as in the present study, photographs of faces are used as stimuli, there may be less tendency to process them by a relatively step-by-step feature analysis. Real faces are unique and not simply permutations of a fixed set of features. Bradshaw and Wallace, themselves, suggest that their implication of serial processing may reflect demand characteristics of their task.

At this point, then, either type of processing appears to be a viable option for the perception of facial expression. The present study sought evidence to discriminate among these options by altering the duration of exposure to a face and measuring the extent to which the face was perceived as smiling. Magnitude estimation of smiles has been shown to be a reliable measure (Simpson & Capetanopoulos, 1968), yielding high test-retest correlations over a 1-month period ($r_s = .64$ to $.86$ for individual Os). In this task, the *gestalt* or simultaneous processing hypothesis predicts that magnitude estimates of smiles will not change with variation in exposure duration. The limiting case would be an extremely short duration where no expression is perceived and magnitude estimates are pure guesses. The serial feature analysis hypothesis predicts that magnitude estimates will change with variation in exposure duration. One specific form of serial processing is represented by an accumulation hypothesis. Suppose O searches a face for, and counts, positive indices of smiling to arrive at his judgment of smile magnitude. Then exposure duration and smile magnitude should interact—judgments of larger smiles should require longer exposure to reach stable or asymptotic levels.

Stimulus Scaling

The logic of the study necessitated a collection of photographs of faces varying in smile magnitude with an accompanying index of the perceived smile magnitude for each face. Front facial views, on 2 x 2 slides, were made of 150 adults attending a summer college session. The purpose of the photographic session was explained to each person and an attempt was made to collect reasonably natural versions of a wide range of smiles. Eight slides (40 males and 40 females) were selected from the set, in an attempt to provide a flat distribution along the continuum of smile magnitude. These were placed haphazardly on a 62 x 21 in. translucent Plexiglas sheet. Each of 18 Os (9 males) then sorted the slides, placing them left to right

*This research was conducted in part with support from NIMH Grant MH-18283-01.

†Now at The Pennsylvania State University.

according to smile magnitude. The Os were asked to position slides according to a ratio scale, keeping distance proportionate to smile magnitude. Time was unrestricted, and Os made frequent rearrangements and adjustments as they worked. Since Os used most, but rarely all, of the available distance for their sorting, there appeared to have been no upper bound restricting their judgments. From these scaling data, final subsets of slides were obtained for the two experiments. Slides selected were those with mean scaled values at one of five equal steps along the smile continuum from "no smile" to maximal smile and with a minimal standard deviation across the 18 Os.

EXPERIMENT I

Method

Forty slides were used as stimuli, four male and four female faces representing each of five smile magnitudes. The slides were rear-projected on opal glass so that a life-size full face was visible with cut-offs in the hair and just below the chin. The room was illuminated so that the projected image of the face was clear but not so bright as to produce any appreciable afterimage. Exposure durations were controlled by a Gerbrands G1166 electronic shutter driven by a Hunter 111C decade interval timer. Four durations were used, .05, .15, .55, and 2.05 sec. One male and one female face of each smile magnitude was presented at each duration. The Os were run in groups of two to four, and specific slide-duration pairings were balanced over groups. A modified magnitude estimation procedure was used. The Os were told to use the number "0" to indicate no smile and the number "1" to indicate the faintest possible smile. Other smiles were to be assigned numbers in proportion to their magnitudes, and hypothetical examples of ratio relationships were given. Six practice slides varying in smile magnitude were presented to provide an example of the range of smiles in use and to assure that instructions were understood. A count-down procedure was followed throughout, so that the brief exposures would not be missed. Slides were presented at a rate of about 4/min, giving ample time to write judgments on an answer sheet. The Os were 40 students (20 males) from introductory psychology classes.

Results

For all Os, magnitude estimates changed in a fashion consistent with the scaled values of the stimuli and were rarely greater than 10. The magnitude estimates were subjected to a repeated-measures analysis of variance. Because of recognized bias in such analyses, all F tests on the

Fig. 1. Mean magnitude estimates made in Experiment I, irrespective of exposure duration. Spacing on abscissa is in terms of scaled values of the stimuli.

repeated measures were made conservatively by reducing the numerator df to 1. Figure 1 presents the mean magnitude estimates of each smile magnitude, irrespective of exposure duration, for male and female Os. These curves can be viewed as discrimination functions, and they reflect slightly greater discrimination among the five smile magnitudes by females than by males [$F(1,1482) = 10.7, p < .01$]. Faces of females were given greater magnitude estimates (3.1 vs 2.9 for males) [$F(1,1482) = 12.1, p < .01$]. This difference was not apparent in the scaled values of the stimuli and could therefore reflect a bias toward perceiving greater smiles on female faces, but the effect is small in magnitude.

The main effect of duration was significant [$F(1,1482) = 8.1, p < .01$], reflecting slightly smaller estimates assigned under progressively longer exposures. The mean estimates were, for the four durations in increasing order, 3.2, 3.1, 2.9, 2.9. Evidence for the accumulation hypothesis should come most directly from a test of the Magnitude by Duration interaction. This interaction was not significant [$F(1,1482) = 3.0$]. In fact, when plotted, the interaction reflected a tendency to assign greater estimates to low-magnitude smiles under the shortest duration of exposure. These observations give little aid to the argument for serial feature analysis. They are clearly inconsistent with the accumulation hypothesis, where exposure duration should have its principal effect upon the larger magnitude smiles and the main effect should thereby show greater estimates as a result of longer exposures.

An alternative way of analyzing the data is by means of correlation. If serial feature analysis is descriptive of the perception of smiles, then correlations of the scaled values of the stimuli, obtained under no time

restriction, vs the magnitude estimates of those stimuli should increase with length of exposure under which the magnitude estimates are made. Thus, median magnitude estimates were calculated for each face under each of the four durations of exposure. For each duration, the medians were correlated with the scale values of the stimuli. The resulting correlation coefficients were, in order of increasing exposure duration, .90, .95, .96, and .94. Neither the magnitudes nor the pattern of these coefficients lend support to the serial hypothesis.

The most striking finding, in both the correlation analysis and the analysis of variance, is the consistency with which the magnitudes of the smiles were perceived, whether the amount of time was extremely short (.05 sec) or unlimited. This is true regardless of the magnitude of smile involved. Whatever information an O needs to judge the smile is obtained very quickly, and additional viewing time adds nothing to alter his judgment.

EXPERIMENT II

There must be some limit of time below which smiles no longer can be perceived. This experiment sought to bracket that limit. If faces are presented too briefly for an O to perceive the expression on the face, but he is nonetheless required to make a magnitude estimate, the discrimination function relating estimates to magnitudes should flatten out as a result of guessing.

A second question raised by this experiment was whether repeated brief exposures to the same face would alter the perceived magnitude of the smile on that face. Thus, some faces were presented three times during the experiment. This also makes possible a second index of guessing. If Os are asked to indicate whether a face is "new" (not seen before) or "old" (seen earlier in the experiment) and the face has been exposed too briefly for perception, then error of identification as old or new should increase. This assumes that, given an exposure sufficient for perception, memory for faces is good, which is now well demonstrated.

Method

The apparatus and procedure were the same as in the previous experiment, with the following exceptions. An 8 x 10 photograph of a female face with an intermediate-magnitude smile was mounted adjacent to the projection screen. The Os were instructed to give this smile a magnitude of 10 and to assign numbers to the test faces relative to this standard. Examples and practice slides were used, as in Experiment I. Fifteen slides were selected from

Experiment I, 3 at each of the five magnitudes, and two duplicates were made of each. These 45 slides were randomly ordered along with a comparable set of 15 additional faces, which each appeared only once in the total set of 60 slides. There were, thus, 30 first-time ("new") presentations of a face and 30 repeated ("old") presentations.

Stimuli were presented at three durations, .02, .12, and 1.02 sec. If a face was repeated within the session, it was always presented at the same duration. The .02-sec duration was the briefest that pilot Os would tolerate under the existing conditions of illumination. Any briefer exposures elicited obvious agitation with the procedure and were reported as unseen. The Os were 24 students from introductory psychology classes (12 males) and were run in groups of 2-6. Specific pairings of slides and exposure durations were balanced over groups. An O responded by writing a magnitude estimate on his answer sheet and then circling "O" or "N" to indicate whether or not the face had been seen earlier in the session.

Results

Those 15 faces which were repeated during the session tended to first occur relatively early in the series. This was not the case for the 15 faces not repeated in the series, which were distributed evenly throughout the 60 trials. Thus, a comparison of the magnitude estimates of those smiles appearing only once with those smiles occurring for the first time (of three) provides a test for any general practice effect. It is necessary to know if judgments of "new" faces change throughout the experiment before it is possible to interpret any change observed upon repeated presentations. An analysis of variance comparing the two sets of slides indicated no difference between the magnitude estimates assigned to them [$F(1,638) < 1$]. There was an indication that the faces in the repeated set had a slightly wider range of smiles [$F(1,638) = 8.11, p < .01$], but this pattern was also evident in the original scale values for the two sets. In general, then, there appeared to be no change in judgment during the experimental session.

The analysis of major interest was performed upon the magnitude

Fig. 2. Mean magnitude estimates of smiles exposed for three durations in Experiment II. Spacing on abscissa is in terms of scaled values of the stimuli.

estimates of the 15 faces which had been repeated during the session. Estimates again closely followed scale values of the smiles. Duration of exposure did interact with smile magnitude [$F(1,968) = 5.59, p < .05$], and duration was also a significant main effect [$F(1,968) = 4.62, p < .05$]. The interaction means are presented in Fig. 2. It is clear that any effect of duration is small, being restricted to an elevation of the magnitude estimates assigned to low-magnitude smiles presented at the shortest duration. This result agrees with the data from Experiment I and disconfirms the accumulation hypothesis. As in the earlier experiment, the outstanding feature of these data is the consistency with which the smiles were judged regardless of exposure duration. Even at the briefest exposure, discrimination among smile magnitudes was excellent.

This pattern of magnitude estimates acquires greater meaning in the context of an analysis of errors of identification of the faces. Table 1 presents the percent error of identification (calling an old stimulus "new," or the converse) as a function of exposure duration and the ordinal number of the presentation. Because

there was an equal number of old and new faces in the series, a chance or guessing rate of error would be .50. Error rate in identifying faces under the .02-sec exposure was not much less than chance (.30), and even an exposure of .12 sec resulted in a relatively high rate of error (.19). There was some tendency toward less error from the first to the third exposure (.22 to .18 to .14). Thus, there was considerable difficulty in recognizing the faces, but in spite of this, the magnitude estimates of the smiles on those faces was nearly as precise as under a long exposure.

The other variables in the experiment had no effect upon magnitude estimation, either as main effects or in interactions. In contrast with Experiment I, discrimination by male Os was as good as that by females. The number of times a face was presented also was immaterial—smiles were judged the same the third time seen as they had been on the first presentation.

DISCUSSION

The results of the two experiments are clear in pointing to the speed at which smiles are differentiated. Very little time is necessary for perception of the magnitude of a smile, and beyond this short exposure (on the order of a few 1/100th of a second?), additional exposure does not appreciably alter judgment. The particular smile in question appears to be quite unimportant. Two interpretations of these results are suggested. One is that the various cues to facial expression, in this case smiling, are processed in parallel—a processing that is nonaccumulative

Table 1
Percent Error of Identification (Chance = .50)

Duration (Sec)	Presentation			Combined
	First	Second	Third	
.02	.36	.30	.25	.30
.12	.19	.21	.18	.19
1.02	.12	.02	.00	.05
Combined	.22	.18	.14	

over time. An alternate interpretation is that, for the O, there is only one stimulus dimension—either only one feature (e.g., the mouth) is attended or the face is seen holistically, as a gestalt. At present, it is impossible to decide between these alternative interpretations. As Garner (1970) has pointed out, before one can logically discuss whether information is being processed serially or in parallel, it must be known that there is indeed more than one informative stimulus dimension. For the perception of facial expression, this is not known. A study by Goldstein & Mackenberg (1966) suggests a way to find out. They selectively degraded familiar faces by covering various portions and found that much of the face was redundant for identification and that the upper portion of the face seemed most informative. Following Goldstein and Mackenberg's procedure, it should be possible to degrade faces in such a way as to discover the locus of information about smiling. As a clue to setting up such a study, it can be noted that the present Os responded to queries about how they had perceived the smiles by indicating the mouth and the eye regions, in about equal numbers, as most informative.

Experiment II showed that even when recognition of a face was poor, perception of the smile on that face was accurate. This might suggest that less information is required for perceiving a smile than for recognizing the face. But again, such a conclusion must await information concerning the locus of information about smile magnitude. Goldstein & Mackenberg (1966) indicated that recognition was carried primarily by the top of the face. If smiling is indicated primarily by the mouth, then both smile perception and face recognition cannot be expected from a single brief exposure, as required in the present study.

Finally, there was some indication in Experiment I that male observers may be less sensitive to smile magnitude than are females. This result was not replicated in Experiment II. But the idea that males are less perceptive of facial expression is not new (see, e.g., Ekman, 1972). However, the sex of an observer seems to make at most a minor contribution to judgment.

REFERENCES

BRADSHAW, J. L., & WALLACE, G.

- Models for the processing and identification of faces. *Perception & Psychophysics*, 1971, 9, 443-448.
- EKMAN, P. Universals and cultural differences in facial expressions of emotion. In J. Cole (Ed.), *Nebraska symposium on motivation*. Lincoln, Neb: University of Nebraska Press, 1972.
- GALPER, R. E., & HOCHBERG, J. Recognition memory for photographs of faces. *American Journal of Psychology*, 1971, 84, 351-354.
- GARNER, W. R. The stimulus in information processing. *American Psychologist*, 1970, 25, 350-358.
- GOLDSTEIN, A. G., & MACKENBERG, E. J. Recognition of human faces from isolated facial features: A developmental study. *Psychonomic Science*, 1966, 6, 149-150.
- HAAF, R. A., & BELL, R. Q. A facial dimension in visual discrimination by human infants. *Child Development*, 1967, 38, 893-899.
- LEWIS, M., KAGAN, J., & KALAFAT, J. Patterns of fixation in the young infant. *Child Development*, 1966, 37, 331-341.
- SIMPSON, W. E., & CAPETANOPOULOS, C. Reliability of smile judgments. *Psychonomic Science*, 1968, 12, 57.
- SMITH, E. E., & NIELSEN, G. D. Representations and retrieval processes in short-term memory: Recognition and recall of faces. *Journal of Experimental Psychology*, 1970, 85, 397-405.
- WOODWORTH, R. S., & SCHLOSBERG, H. *Experimental psychology*. New York: Holt, Rinehart & Winston, 1954.
- YIN, R. K. Looking at upside-down faces. *Journal of Experimental Psychology*, 1969, 81, 141-145.