Information content and the identification of human faces

SUSAN L. MITCHELL, ROBERT PASNAK, and JANICE W. CAMPBELL
George Mason University, Fairfax, Virginia

The identifiability of men's faces was predicted from their information contents. A sample of 12 photographs of college men's faces with similar hair colors, eye colors, and complexions was obtained. Student judges, who received systematically varied instructions, determined the information contents of these faces and their interior features by counting the number of times their contours changed direction (e.g., made an angle or arc). Other students tried independently to identify the faces when they were presented in a tachistoscope for durations of 300, 600, or 900 msec. Errors were negatively correlated with the information contents of the faces, especially the information content of the mouth. Presentation times and instruction effects were also observed.

In this paper we explore a new basis for predicting the role of a face's internal features in identification of the face, and for predicting the relative efficiency with which different faces can thereby be identified. The starting point is Attneave's (1954) classic definition of information content, which may be applicable to the features of the face. Attneave showed that the information content of any shape could be given by the number of points it would take to make an outline of that shape. The points in question were those at which the contour changed direction—that is, where it made an arc or an angle. Attneave defined each such point as a "bit" of information. According to this definition, a highly irregular contour that changed direction many times would contain many bits of information. A contour that was smooth and regular—seldom changing direction—would contain, and therefore could be defined by, less information.

Attneave went on to show that subjects were in general agreement as to which points of a contour constituted the "bits" of information. Thus, when selecting the points needed to define a randomly generated potato-shaped contour, subjects working in isolation and without feedback showed a marked tendency to pick the same points. This may indicate that points of maximum curvature have some perceptual saliency.

Attneave (1954) further offered a demonstration that an image composed of such points is recognizable. He took the form of a sleeping cat, subjectively selected the points of maximum curvature, and then connected those points with straight lines. The result was easily recognized as an image of a sleeping cat, although he only used 38 points.

Seman, Pasnak, and Tyer (1976) built upon Attneave's work by showing that the information contents of photos

of human hands could be so quantified. They tested whether this quantification was related to discrimination of the photos. In their paradigm, a negative correlation ($r = -.53$) between information content and discrimination was found, but only for the region of the thumb.

The human face is much more complex and variable than the human hand. The present study represents an effort to quantify the face and its features according to the method suggested by Attneave, and to test the quantification's effectiveness in predicting errors in identification of the face.

METHOD

Subjects

No subject participated in more than one phase of the study, and all were treated in accordance with ethical principles of the American Psychological Association (APA, 1981). Models for the photographs were 38 white males from introductory psychology classes, who posed in a chinrest with mouths closed, an unsmiling expression, and eyes fixated on the camera. The 12 most similar photographs were chosen for use in the experiment. The remaining participants (judges and experimental subjects), who were from advanced classes in child development, averred they neither knew nor had seen any of the 12 models in the photographs.

Apparatus

A Canon AE-1 Program 35-mm camera with a Vivitar Series 1 70-210 mm lens was used to take the photographs on Kodak Tri-X pan black-and-white film. The 12 selected photographs were mounted on black posterboard and labeled A through L. Ten different random orders of mounting were used in the course of the experiment.

Duplicate photographs were uniformly trimmed ¼ in. above the eyebrows, down the side of the face 1½ in., diagonally across the cheeks to ¼ in. from the corners of the mouth, and then straight across ¼ in. below the bottom lip. This was done to eliminate face size, hair texture, and hair outline as variables. The photographs were then mounted on a black background to fit inside a Lafayette Instruments Model 42015 tachistoscope, the interior of which was also black.

Procedure

Two experimenters and 24 student judges scored the faces for information content. This involved counting the number of points necessary to define the contours of the eyes, eyebrows, nose, and mouth. The students were given an unlimited amount of time to score the photographs.

Extensive pilot work (Pasnak, 1969; Mitchell, 1985) indicated that instructions had an important effect on this quantification; therefore, different types of instructions were given at two sessions. The following "geometric" instructions were given to eight judges who attended one session.

> You are to determine the number of dots it would take to completely outline each feature—eyebrows, eyes, nose and mouth—of these faces. To do this, a dot would be needed at each point where the contour changes direction. If you placed a dot at every such point, you would have a good connect-the-dot type of outline of each face, much as you might see in a child's workbook.
> You'll find this task difficult, since it is hard to be sure where the contours of the features change directions. But if you are diligent, you should be able to do a reasonable job. Please count every change of direction of contour of the features, and include their inner parts as well as their outlines.

The 16 judges who attended the other session were given "phenomenal" instructions. These were identical to the geometric instructions, except for the last sentence. This was changed to read as follows: "Please count the changes of direction of contour you think important for each feature, and include their inner parts as well as their outlines."

The experimenter demonstrated the desired performance to each group of judges with a sample photograph not used elsewhere in the experiment. For the first group, the experimenter meticulously indicated every point at which the contour changed direction. For the second group, minor changes or details were ignored. Each judge then assessed the photographs, which were presented individually in arbitrary, variable orders, and tallied his or her results on individual score sheets. The first and second authors also scored the photographs, separately from the judges and from each other, following the geometric approach.

The identification task involved showing 19 experimental subjects the trimmed photographs individually in the tachistoscope and asking them to choose which one they believed they were shown from a display of all 12 photographs in untrimmed form. The trimmed photographs were shown once each at 300, 600, and 900 msec, according to random schedules. After each presentation, the subjects were given unlimited time to select the photograph they had just seen from the 12 displayed on the posterboard. The subjects were required to choose even if they were uncertain; only after having made a choice could a subject trigger the tachistoscope for the next trial. Most subjects took 20-60 sec between trials. Note that individuals trying to identify or quantify the faces were blind to all efforts but their own.

RESULTS

The errors made in identifying each face are given in Table 1. A directional F test showed that identification

Table 1
Identification Errors at Three Presentation Times

Face	Time (in milliseconds)			
	300	600	900	All
A	3	2	4	9
B	4	5	3	12
C	6	6	9	21
D	8	8	6	22
E	7	5	5	17
F	12	8	7	27
G	12	6	0	18
H	10	7	5	22
I	8	8	4	20
J	11	6	7	24
K	6	7	6	19
L	10	5	6	21

Table 2
Information Content Obtained by Geometric and Phenomenal Scoring

Judges	Brows	Eyes	Nose	Lips	Sum
		Geometric Scorers			
1	47.3	53.5	29.5	30.0	160.3
2	39.6	47.2	32.7	28.8	147.8
3	44.3	46.3	29.2	26.7	146.3
4	29.4	41.7	20.6	21.8	112.7
5	49.3	42.2	31.7	32.9	156.7
6	28.3	58.2	25.9	23.8	135.3
7	43.7	45.2	25.9	20.2	133.2
8	37.0	48.0	18.5	23.7	127.3
S.M.	30.0	43.8	24.4	20.0	118.7
R.P.	33.2	60.4	24.5	26.5	147.4
		Phenomenal Scorers			
1	14.2	29.2	17.6	16.5	76.7
2	20.9	39.9	18.9	20.5	100.2
3	18.7	37.8	16.3	15.3	88.2
4	23.3	41.6	27.3	25.6	117.8
5	20.2	28.4	17.3	19.2	85.2
6	19.4	29.7	15.1	16.2	80.5
7	14.7	27.9	15.2	12.2	70.0
8	15.4	35.4	18.2	14.9	83.9
9	39.7	46.2	23.8	24.4	134.2
10	32.1	37.0	15.4	15.6	100.1
11	16.9	33.7	12.4	10.1	73.1
12	10.2	20.3	11.9	10.4	51.2
13	19.2	31.4	21.1	16.1	87.8
14	18.0	27.9	16.4	11.2	73.5
15	21.2	38.2	16.3	17.2	93.0
16	16.7	36.7	20.8	10.9	81.9

Note—Means given are for the 12 faces.

errors decreased as presentation time increased [$F(2,22) = 3.10, p < .05$]. This is similar to the finding of Ellis, Davies, and Shepherd (1977) for unfamiliar faces. The distribution of total errors for the 12 faces is approximately normal, as would be expected for such arbitrarily selected stimuli.

The average number of contour changes (information content) for the features of the faces as determined by geometric or phenomenal scoring is given in Table 2.

Pearson product-moment correlation coefficients are the most appropriate statistic with which to describe and assess the relation between the number of contour changes (information content) determined for the faces and their features and the number of identification errors made. For 12 stimuli, a linear prediction from information content would have to account for more than 33% of the variation in number of identification errors between faces to attain statistical significance (i.e., $r > .576$).

When the faces were scored geometrically, the relation between their total information content and the number of errors made in identifying them was negative (see Table 3). The counts of the number of contour changes made by the senior author and those by one of the judges correlated significantly with identification errors across all presentation times. When the results for the different presentation times were considered separately, a some-

Table 3
Correlations Between Information Content and Errors
at Different Presentation Times

	Phenomenal Scoring			Geometric Scoring		
	Correlation	Judge	PT	Correlation	Judge	PT
Brows	−.58	3	300	−.58	6	600
	−.60	7	300			
	−.60	11	300			
Eyes	.62	3	300	(none)		
	.62	7	600			
Nose	−.72†	12	300	−.72†	8	900
	−.76†	12	600	−.62	S.M.	900
	.58	14	600	−.63	S.M.	all
	−.64	6	900			
	−.81†	12	all			
Mouth	−.60	1	300	−.62	1	600
	−.68*·	2	300	−.58	1	900
	−.68*	2	all	−.74†	5	all
	−.60	7	600	−.60	S.M.	all
	−.66*	11	300			
	−.58	11	all			
	−.64	12	300			
	−.72†	12	600			
	−.66*	12	all			
	−.81†	16	300			
	−.64	16	900			
Whole Face	−.62	3	300	−.72†	S.M.	300
	−.72†	12	600	−.62	S.M.	600
	.58	14	600	−.86†	S.M.	all

Note—Only significant ($p < .05$) Pearson product-moment correlations are given. *$p < .02$.   †$p < .01$.

what different picture emerged. At 900 msec, identification errors were significantly related to counts made by 3 of the 8 judges, and by the senior author. The strength of the relations gradually decreased as presentation time decreased, so that at 300 msec none of the judges' counts produced significant correlations.

Of the 16 judges instructed to take a phenomenal approach to scoring the information content of the faces—counting only those changes of contour that seemed subjectively significant—all but 5 produced counts that were significantly correlated with identification errors at one presentation time or another. In contrast to the results obtained with geometric instructions, most of the significant relations emerged at the shortest presentation time (see Table 3), a glance so short as to limit eye movements.

The majority of the significant correlations were for the mouth. The eyes were the feature whose information content was least likely to be related to identification errors, and the nose and brows were intermediate. Some judges were much better at producing predictive quantifications of some features than others.

Two of the correlations obtained with phenomenal scoring were positive. These were the quantifications of Judge 14 for the brows and for the total face. All the other significant correlations were negative, leading to the conclusion that these two correlations most probably represent Type I errors.

A few such errors must be expected when many independent correlations are calculated. The consistent pattern of negative correlations, many associated with very low alpha (p) levels, indicates that the more general finding of an inverse relation between identification and information content should not be dismissed on these grounds.

DISCUSSION

This research obviously represents only a beginning at assessing the information content of faces. Much of the behavior of judges and identifiers has not yet been brought under stimulus control. Yet this method of quantifying faces shows some promise, and some of the problems can now be delineated.

Although many of the judges' efforts to quantify the faces were not related to identification errors, many others produced significant results, and the magnitudes of the relations found were not small. The subtlety of the task accounted for some of this variation in results. The contours were irregular curves of varying brightness gradients, and there were unavoidable uncertainties as to when a contour had really changed direction. Thus the second experimenter counted 25% more changes, on the average, than did the first. In addition, the students varied widely in their seriousness and patience, even if they understood their task perfectly. Such factors must have had an impact, and they probably accounted for many of the failures to obtain significant correlations.

Furthermore, narrowing the pool of stimuli to the 12 most similar faces decreased the power of the analysis so that only relatively strong correlations could be statistically significant. Although this may have been an overly stringent approach, it is such cases—involving discrimination between similar stimuli—that provide the best test of the stimulus quantification. Another statistical limitation was that the correlation coefficient reflected only the straight-line component of relations between variables, whereas nearly all relations were somewhat curvilinear. Even with these difficulties, more than half of the judges produced quantifications that for at least some part of the faces as some presentation time predicted other students' perceptual errors with a reasonably high degree of success. We were led to conclude that this measure of contour complexity or information content must have reflected a stimulus dimension that has an important role in face perception.

The importance of the quantification varied with presentation time. The geometrically obtained scores were the ones most likely to predict errors at longer presentation times. This suggests that attention to or processing of the small complexities reflected by these counts could not proceed immediately. Given time, however, the viewers did make decisions about the identity of the faces viewed, which corresponded to a few of these relatively detailed quantifications of some facial features.

The phenomena scorings were the ones most likely to predict errors successfully at very short presentation times. This suggests that the most relevant contour changes, as distinguished by these scorers, were attended to and processed early. Since the relation declined as presentation time became longer, other components of the "facial episode" (Klatzky & Forrest, 1984) must have come into play and diluted the effect of information content on stimulus identification. Given the number of factors involved—part-whole and spatial relations, abstract visual properties, verbally encoded information, semantic and affective attributes, and attentional, processing, and memory variables, to name but a few—the potential interactions with information content were quite complex, and its effects must often have been submerged by those of other variables. Thus, the role of information content in face identification is necessarily limited. This preliminary research does, however, suggest that it is of some importance. If it can be made a manipulable variable, efforts to assess its impact might prove quite fruitful.

REFERENCES

AMERICAN PSYCHOLOGICAL ASSOCIATION. (1981). Ethical principles of psychologists (rev. ed.). *American Psychologist*, **36**, 633-638.

ATTNEAVE, F. (1954). Some informational aspects of visual percep-
tions. *Psychological Review*, **61**, 183-193.

ELLIS, H. D., DAVIES, G., & SHEPHERD, J. W. (1977). Experimental
studies of face identification. *National Journal of Criminal Defenses*,
3, 219-234.

KLATZKY, R. L., & FORREST, F. H. (1984). Recognizing familiar and
unfamiliar faces. *Memory & Cognition*, **12**, 60-70.

MITCHELL, S. L. (1985, April). *Information content as a determinant
of the identification of human faces*. Paper presented at the meeting
of the Virginia Psychological Association, Richmond, VA.

PASNAK, R. (1969, April). *A quantitative method for predicting the dis-
criminability of human faces*. Paper presented at the meeting of the
Eastern Psychological Association, Philadelphia, PA.

SEMAN, W. A., PASNAK, R., & TYER, Z. E. (1976). Discriminability
and identifiability of the human hand based on its quantified com-
plexity. *Perceptual & Motor Skills*, **42**, 287-293.

Announcement

Conference on Flashbulb Memories and the Shuttle Explosion
Atlanta, Georgia
February 2-3, 1990

Call for Papers

On February 2-3, 1990, the Emory Cognition Project will host a working conference on "Flashbulb" Memories, with a special focus on memories of the 1986 explosion of the space shuttle *Challenger*. This conference will be jointly sponsored by the Office of Naval Research and the Air Force Office of Scientific Research.

We are especially interested in empirical studies of people's ability to recall how they first heard the news of this disaster. If you have data relevant to this issue, published or unpublished, please call or write: Eugene Winograd (404-727-7448) or Ulric Neisser (404-727-7973), Department of Psychology, Emory University, Atlanta, GA 30322.