Journal of Applied Psychology
1980, Vol. 65, No. 3, 307-316

Sketch Artist and Identi-Kit Procedures for Recalling Faces

Kenneth R. Laughery and Richard H. Fowler
University of Houston

Sketch artists and the Identi-kit provide face construction techniques widely employed in law enforcement. The effectiveness of these techniques was explored in a study in which 142 subjects worked with artists or Identi-kit technicians to construct from description a sketch or an Identi-kit composite for each of 71 different white-male target faces. The artists and technicians also prepared a sketch and composite while directly viewing each target face. Ratings of goodness of fit between the sketches/composites and photographs indicated that sketches were superior to composites. Artist differences were found, but technician differences were minimal. Sketches from view were better than from description, but the description-view variable did not affect composites. These latter two results indicate that the Identi-kit technique may have serious limits in representation accuracy. Time-line analyses of work on various features revealed that subjects "move around" more and take longer in constructing sketches. Results are discussed in terms of the utility of these and other face construction procedures.

Recent psychological research on memory for faces has usually employed recognition procedures. The reason for using recognition tasks, of course, is the response problem—Most people are not capable of producing (drawing) a facial image that would accurately reflect their memory for the face.

An exception to the use of recognition tasks is a series of experiments by Ellis, Shepherd, and Davies (1975) and Ellis, Davies, and Shepherd (1978). These studies explored the use of the Photofit technique for recalling faces. The Photofit system consists of numerous alternatives of the following five facial features arranged in a booklet: forehead and hair, eyes, nose, mouth, and chin. These features are black and white photographic prints taken from pictures of real faces. Witnesses examine the features and select those closest to the face they are trying to reconstruct. The selected features are placed together to make the face, which can then be revised. The overall conclusion drawn from these experiments is that the Photofit system does not lead to good facial representations and that the limitation is primarily in the design of the system itself.

These results are disappointing in two respects. First, the utility of the Photofit as a law-enforcement procedure is obviously limited by the extent to which it leads to accurate representations. Second, the ability to employ recall procedures in studying memory for complex visual configurations, such as faces, would be a valuable tool in furthering our understanding of human memory. The Photofit

This project was supported by Grant 76-NI-99-012 awarded by the Law Enforcement Assistance Administration, U.S. Department of Justice, under the Omnibus Crime Control and Safe Streets Act of 1968, as amended. Points of view or opinions stated in this document are those of the authors and do not necessarily represent the official position or policies of the U.S. Department of Justice.

Ben T. Rhodes made significant contributions to the planning and implementation of this project. Glen Duval, Mike Mauldin, and Sharon Neyland contributed to the often difficult tasks of recruiting and coordinating subjects, compiling and analyzing data, and the countless other details required in a complex research effort. To them and others, our thanks.

Requests for reprints should be sent to Kenneth R. Laughery, Department of Psychology, University of Houston, Houston, Texas 77004.

appears to be of limited value in this regard.

Two other widely used facial construction techniques in the law-enforcement field are provided by the sketch artist and the Identi-kit. The sketch artist procedure involves an artist sketching the target person while getting information from a witness through conversational interaction. The Identi-kit is a set of transparent celluloid sheets, each containing a line drawing of a facial feature. There are a large number of sheets for each feature (i.e., many types of noses, eyes, etc.). A trained technician constructs a composite face by interacting with a witness to select appropriate features that are then superimposed to make a face. A special marking pencil is available for the technician to make additional modifications or to add detail.

The study reported here was undertaken to explore the utility of the sketch artist and Identi-kit techniques for constructing facial images from memory. The investigation is worthwhile in terms of applied, forensic uses as well as in potential uses for investigating facial recall, as opposed to recognition.

There was one hypothesis and several other points of interest in the study. The hypothesis was that sketches would be better representations than Identi-kit composites for two reasons. First, the Identi-kit consists of a finite set of features, whereas the sketch artist can presumably generate an infinite set. With the Identi-kit there may be occasions when the "right nose" simply is not there. A second reason for possible sketch superiority is that some kinds of detail, such as shading, are typically added in sketches but less available with the kit. Previous work by Davies, Ellis, and Shepherd (1978) has shown that such detail enhances recognition. It should be noted that there is a possible reason for predicting the opposite outcome. Since witnesses may have difficulty communicating about facial features, the Identi-kit could be better, since the witness may be able to recognize an appropriate feature representation from the set of available alternatives.

One point of interest concerned two procedures that were used. In one case the artist/technician constructed the image from a witness's description (standard law-enforcement procedure), whereas in the other case the image was produced while the artist/technician viewed the target. Ellis et al. (1975) reported significantly better Photofit constructions when the image was done with the target face in view as opposed to a memory condition. However, in another Photofit study Ellis et al. (1978) did not find a significant effect of this variable. Another point of interest concerned artist/technician effects on the image generation outcome. Put simply, it was anticipated that some amount of variation in image quality (goodness of fit) would result from different people serving as artists/technicians. These two points are potentially important, since the presence or absence of view/description and artist/technician effects would have implications for the locus of technique limitations. Specifically, no effect of these variables would imply that the major limiting factor in the quality of images is the technique itself, not the skills of the artists or technicians.

In addition to the previously mentioned hypothesis and issues, several other aspects of the facial image generation task were explored. Questions regarding relationships between witness characteristics and image quality were examined using correlational procedures. For example, imagery and verbal abilities were correlated with image quality. One reason for being interested in these relationships is the possibility of distinguishing between good and poor witnesses. If reasonably straightforward techniques were available for assessing a witness's ability, and if these measures correlated with image quality, one would be in a position to put more or less confidence in an image produced by a particular witness. Similarly, if strong correlations exist, further research might be appropriate for improving the quality of images produced by witnesses expected to do poorly.

Method

The study consisted of two phases, construction and rating. The construction phase dealt with the image generation part of the study; that is, subjects saw a target person and then worked with an artist or

technician to generate a facial image. The rating phase was an experiment in which the images generated in the first phase were evaluated for goodness of fit by a separate group of subjects. Procedures for these two phases will be described separately.

Construction Phase

This phase consisted of the image generation part of the overall study.

Subjects. The subjects can be divided into two groups, those who served as targets and those who served as witnesses. A total of 71 target subjects were used, all white males. The targets were drawn from the student body at the University of Houston and from the Houston community at large. The only restriction placed on the selection of targets, beside being white males, was that they be unknown to the witness subjects, the sketch artists, and the Identi-kit technicians. There were 142 witnesses with no restrictions placed on their selection. Most witnesses were students at the University of Houston. A breakdown of these subjects by sex and race shows 41 males (37 white, 1 black, 2 Chicano, and 1 Oriental) and 101 females (81 white, 9 black, 8 Chicano, and 3 Oriental). All subjects were volunteers and were paid $2 per hour for participating.

Design. The design included three experimental variables. The first was the image generation technique, consisting of the sketch artist or the Identi-kit. The second variable, to be referred to as artist/technician, consisted of three artists and three Identi-kit technicians. This variable was nested within technique; that is, the three artists and the three technicians were six different people.

The third variable was created by an auxiliary procedure in the experiment. After images had been completed from the witnesses' descriptions, the artist and technician each constructed a second image while viewing the target directly. Thus, for each target and technique condition there were two images, one from description and one from view. This variable is referred to as target presentation.

Artists and technicians. Three people served as sketch artists and three others as Identi-kit technicians. The three artists, SN (female), BM (male), and AM (male), were recent graduates of the University of Houston with a Bachelor of Fine Arts degree. All three had a good deal of training and experience in portrait work. Prior to beginning the experiment, each artist constructed several practice images from description. The three technicians, MM (male), RF (male), and JH (female), were enrolled as graduate students in psychology at the University of Houston. MM attended a 2½-day Identi-kit training course sponsored by the Identi-kit Company. RF and JH were trained in the procedures by MM. All three technicians practiced extensively prior to the experiment.

Procedure. The experimental task consisted of two parts, target exposure and image generation. During the first part the target was exposed to two witnesses. The image generation followed and consisted of one witness working with a sketch artist and the other witness working with an Identi-kit technician.

The procedural aspects of each experimental session involved the following six people: the experimenter, a sketch artist, an Identi-kit technician, a target, and two witnesses. Since it was necessary to carefully control the timing and manner in which different individuals encountered each other, and because a variety of data was obtained from the various individuals, a relatively complex and carefully controlled procedure was carried out. Details of the procedure are available elsewhere (Laughery, Duval, & Fowler, Note 1). The following description provides an outline.

Two witnesses reported to a room where they filled out a subject data form. This form asked for information about the witness, including certain physical characteristics. Instructions were then presented to the witnesses including a description of the target exposure and image generation parts of the study. The target meanwhile reported to an adjacent room where the experimenter, after finishing with the witnesses, instructed the target on the nature of the study.

Following the instructions the witnesses were escorted to the room where the target was located. With everyone seated at a table, the experimenter moderated an 8-minute conversation, which is referred to as the *exposure period*. To the extent possible, the discussion focused on the target's interests, activities, and so on. Although the setting may seem to have been somewhat strained or artificial, in actual practice it generally proceeded smoothly with reasonably good conversation.

Following the exposure period one witness was escorted to a room to work with a sketch artist to generate an image while the other witness went to work with an Identi-kit technician. At the beginning of the image generation phase, each witness filled out a general description form about the target that was used by the artist/technician as a starting point. Then the witness and artist/technician interacted to construct the image. The verbal interaction was tape recorded.

Following the exposure period, the target completed the subject data form and then posed for a bust-length front photograph. After completing the images, witnesses filled out three additional forms. The first was a subject comments sheet that solicited comments regarding the manner in which they carried out the task. The other forms were the Betts Vividness of Imagery Scale and Gordon Test of Visual Imagery Control (Richardson, 1969); both are paper-and-pencil tests of imagery or verbal memory. The final step consisted of the artist and technician producing a sketch and composite while viewing the target directly.

Each artist and technician completed images for 24 targets, except SN and MM who did 23.

Rating Phase

An important set of issues in this study concerns the manner in which one evaluates facial images. What does one measure? How does one decide whether a particular image is a good, fair, or poor representation of a real face, and how is this goodness of fit quantified? Ellis et al. (1975, 1978) used various rating procedures in which judges simultaneously viewed the target face and the image and rated them for goodness of fit. A

similar approach was used here. A rating procedure was employed in which an independent group of subjects rated each image–photograph pair for goodness of fit on a 6-point scale.

Subjects. Sixty-four students enrolled in an introductory psychology course at the University of Houston served as subjects. Extra credit was given in the course for participating.

Task. The task consisted of showing the subjects a sequence of pairs of slides. Each pair consisted of a target photograph and one of the four images for that target. The pair was projected onto a screen in front of the subject for 10 sec. The projected images were approximately life-size. The subject looked at the images, made a decision regarding the goodness of fit of the image to the photograph, and then indicated the rating on a response sheet. The ratings were made on a 6-point scale, where the two ends of the scale were defined as "most similar" and "least similar."

Design. The similarity ratings were collected from two different groups of subjects, each group rating a different subset of the 71 faces. The reasons for this procedure were twofold, both logistical. First, the image generation experiment was spread over a long time period, and it was desirable to complete some analyses before the entire data collection phase ended. Second, the number of ratings to be provided by each subject was four times the number of target faces. Therefore, if all 71 were introduced in one session, subjects would be required to complete 284 ratings. Such a procedure potentially introduces factors such as fatigue. An analysis of the task led to the conclusion that about 200 ratings is a reasonable maximum. As a result, ratings were obtained on 51 targets (204 images) in the first part of the experiment (51 instead of 50 was simply a convenience due to the availability of stimulus materials). The other 20 targets (80 images) were rated in the second part.

Both parts of the experiment consisted of a rather elaborate ordering and counterbalancing of the sequence of images across different subjects. The purpose of introducing this procedure was twofold. The first related to the fact that for each different target face, four images were to be rated. These images were the sketch from description, sketch from view, composite from description, and composite from view. It was important that the image slides for a particular target face not appear too close together because each rating should be independent of how well the other images matched that target. The second reason for counterbalancing was to eliminate practice effects. Twenty-four subjects were run in the first part of the experiment, and 40 were run in the second.

Materials. The materials consisted of 355 slides. These included a photograph and one of each of the four different images for each of the 71 targets.

Procedure. Subjects were brought into the laboratory where they sat in a classroom-type desk. The viewing screen was located approximately 3 m in front of the subjects, and two Kodak Carousel projectors were above and behind them. Instructions were read informally, and the subjects were given a set of response sheets. A series of 10 sample pairs were then presented to familiarize subjects with the task. This sample included pairs representing a range of goodness-of-fit values as determined in pilot work. The entire set of pairs (204 or 80) was then presented at a 10-sec rate. In all pairs the photograph appeared on the left and the image on the right.

Results

Several analyses were carried out on the data. The ratings were quantified 1–6, in which 1 was a very good ("most similar") fit and 6 was a very poor ("least similar") fit. Figure 1 shows an example of a good and poor sketch and composite. Using this rating as a dependent measure, an analysis of variance examined the effects of four variables: replication (the two parts of the rating experiment), technique (sketch artist or Identi-kit), artist/technician (which was nested within technique), and presentation (description or view). The mean rating for each of the conditions is shown in Table 1.

The main effects of three variables were significant: technique, $F(1, 46) = 134.24$, $p < .001$; presentation, $F(1, 46) = 174.23$, $p < .001$; and artist/technician, $F(4, 184) = 19.54$, $p < .001$. Images were better if done as sketches or from view. The artist/technician effect indicates simply that there were differences between artists and technicians. The means for the three artists were 3.1, 2.9, and 3.4 for SN, BM, and AM, respectively. The Identi-kit technician means were 4.1, 4.0, and 3.9 for RF, MM, and JH. It appears that artist differences may be significantly greater than technician differences; however, this interaction cannot be statistically examined in the present study, since the artist/technician variable is nested within technique.

The Technique × Presentation interaction was significant, $F(1, 46) = 68.59$, $p < .001$. The data indicate a larger difference between view and description in the sketches (2.7 vs. 3.6) than in the composites (3.9 vs. 4.1). A significant Presentation × Artist/Technician interaction, $F(4, 184) = 13.85$, $p < .001$, simply reflects larger differences between view and description for some artist/technicians than others.

Although the replication variable did not produce a significant main effect, it did interact with technique, $F(1, 46) = 15.19$, $p < .001$, and presentation, $F(1, 46) = 28.43$, $p < .001$. The effects of technique

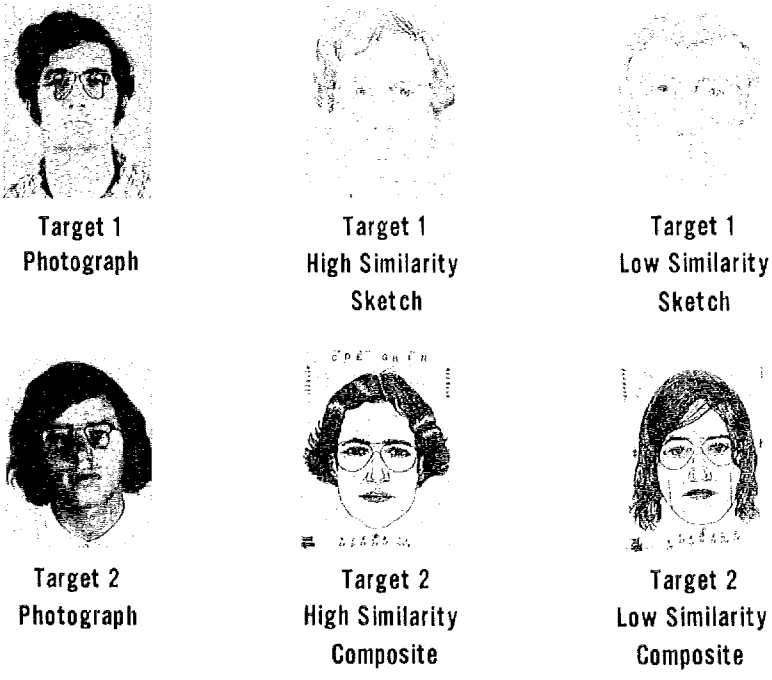

Figure 1. Examples of good and poor sketches and composites.

and presentation were in the same direction but greater in the second replication.

Although it was not part of the formal analysis of the rating experiment, an informal aspect of the outcome is worth mentioning. It was common for subjects who had completed the rating phase to comment on how poor the overall quality was. This general impression of low-quality representations is consistent with the conclusions of Ellis et al. (1975, 1978) on the Photofit technique.

As noted earlier, a variety of data was obtained in addition to the images. Included in the data were scores on the Betts and Gordon imagery tests for witnesses. Also, for subjects who were students at the University of Houston, Scholastic Aptitude Test (SAT) Verbal scores were obtained.

Six correlations were computed: The

Table 1
Means and Standard Deviations of Ratings on 1–6 Similarity Scale

		Sketch						Identi-kit					
		SN		BM		AM		RF		MM		JH	
Rating study	Presentation	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Replication 1	Description	3.5	1.1	3.5	1.3	3.6	1.3	3.9	1.5	3.9	1.4	3.8	1.4
(51 targets)	View	2.7	1.1	2.7	1.2	3.4	1.1	3.9	1.3	3.7	1.5	3.8	1.3
Replication 2	Description	3.7	1.0	3.5	.9	3.7	1.0	4.4	1.2	4.6	1.0	4.2	1.0
(20 targets)	View	2.3	1.1	2.0	1.1	2.8	1.0	4.1	1.3	3.9	1.2	3.8	1.1

Note. Lower scores represent better images. The letters SN, BM, AM, RF, MM, and JH refer to the initials of the artists and technicians.

Table 2
Means and Standard Deviations of Time-Line Measures

	Different feature codes		No. of feature stops		Time per feature stop (sec)		Total time (min)	
Technique	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Sketch	13.2	2.5	30.0	11.7	70.9	16.5	34.24	9.0
Identi-kit	7.9	2.3	11.6	5.4	112.7	54.6	21.69	10.8

ratings for sketches from description and composites from description were each correlated with the two imagery scores and the SAT Verbal scores. Two of the outcomes were significant: The correlation between the sketches and the Gordon imagery score ($r = -.213, p < .05$) and the correlation between the composites and the SAT Verbal scores ($r = -.487, p < .01$). These results are in the expected direction; that is, better images related to greater imagery and verbal ability. (The lower the rating, the better the image, hence the negative values.) None of the other correlations approached significance.

Correlations were computed between the ratings of the two types of images and the total time used to generate the images. The latter measure was defined from the beginning of the subject's interaction with the artist/technician until the subject stated that the image was as good as he/she could construct (or words to that effect). Both correlations had a value of .06, which was not significant.

The sex and race characteristics of the witness subjects served as a basis for analysis. The ratings for images done by male and female witnesses were compared as were the images generated by the different races. None of these results were significant.

In addition to the goodness-of-fit and correlational analyses, several aspects of the results were examined by what is herein referred to as time-line analysis. During the actual process of generating the images, tape recordings were made of many of the verbal interactions between the artists/technicians and witnesses. The tapes of 62 interactions were available for detailed analysis. Twenty-three features were defined on the basis of the contents of the tapes and the experience of the artists/technicians. The 23 features were eyes, nose, mouth and lips, ears, forehead, cheeks and cheek bones, jaw and jawline, chin, hair, hairline, eyebrows, sideburns, moustache, beard, face shape, proportions, glasses, eye color, complexion, wrinkles and face lines, general expression, scars and moles, and neck. These features represent a fine-grained breakdown of the face. Such fine detail is appropriate in developing a first-stage classification scheme, since it is a relatively simple matter to combine features later.

Following the definition of the 23 different feature codes, the boundaries between work on each successive feature were identified on the tapes. A feature stop is defined as the continuous work on a given feature. It should be noted that the number of feature stops will exceed the number of feature codes, since witnesses typically work on a given feature code more than once. The last step in analyzing the tapes was to note the time lapse for each successive feature stop. To summarize, the output of this analysis was the sequence in which the features were worked on and the length of time spent on each.

Means and standard deviations for the different measures by technique are shown in Table 2. The technique differences are clear. In creating sketches, witnesses used a greater number of feature codes, made more feature stops, spent less time per feature stop, and used more total time.

A second analysis of the time-line data focused on the different features. The proportion of feature stops to total time was computed. These measures reflect the relative amounts of time and effort devoted

Table 3
Most Attended Features in Time-Line Feature Analysis

Proportion of feature stops to total stops				Proportion of feature time to total time			
Sketches		Identi-kit		Sketches		Identi-kit	
Feature	Proportion to total	Feature	Proportion to total	Feature	Proportion to total	Feature	Proportion to total
Hair	.140	Hair	.151	Eyes	.177	Hair	.193
Eyes	.117	Nose	.119	Hair	.174	Eyes	.186
Face shape	.091	Eyes	.113	Nose	.126	Nose	.149
Chin	.091	Eyebrows	.105	Mouth & lips	.072	Eyebrows	.108
Nose	.084	Chin	.097	Chin	.072	Mouth & lips	.088

to the various features. Table 3 shows the five features that received the most attention for each technique. Clearly, there is a great deal of consistency across techniques in how much time and effort is devoted to the various features.

A factor of potential importance in generating facial images is the experience of the artists/technicians. A brief description of the training and experience of each artist/technician was presented earlier. To examine possible learning effects, an analysis of the quality of images as a function of experience was carried out. Each of the artist's/technician's images were grouped into blocks of five, and an analysis of variance was carried out on the ratings across blocks. There was no indication of any improvement or decline for any artist or technician.

It seems likely that faces vary in terms of the ease or difficulty with which accurate representations can be created for them. If the sketches and composites prepared from view represent the "best possible" images, and if the images from view and description are positively correlated, then it could be concluded that one limiting factor in image quality is related to characteristics of the target face. For sketches the correlation between images from view and description was significant ($r = .334, p < .01$). Similarly, a significant correlation was found for composites ($r = .363, p < .01$).

Discussion

The results indicate that sketch artists produce better images than the Identi-kit. There are several factors that could account for the superiority of sketches. First, there is a limited set of alternative faces one can create with the Identi-kit, whereas a sketch artist can produce an essentially infinite set. Hence, with the Identi-kit there may be times (and according to technicians, there are) when "the right nose is not there." A second reason for sketch superiority may be the additional detail such as shading, age lines, and so forth, that typically is more predominant in sketches than in composites. This added detail is reflected in the time-line data where a greater number of feature codes are used for sketches than for composites. A related fact here is that a recent version of the Identi-kit (updated since this work was completed) includes a wider range of features and greater shading —changes that should lead to better representations.

A third possibility may be related to the total time difference between techniques. More time is spent generating sketches than composites. More time is not directly the point, however, since the time difference could be accounted for simply by the fact that an artist requires more time to produce a feature than the Identi-kit, with which features are simply selected. The key point is that because of the greater production-time requirements of the sketch, the witness spends more time thinking about the target, which may lead to a more accurate memory and description. There is a serious hitch in this explanation, however, since the correlational results showed that within techniques, total time was not related to goodness of fit. A fourth possible explanation

emerges from the time-line data. In generating sketches, witnesses use more codes, make more feature stops, and spend less time per feature stop. These differences seem to reflect more "moving around" in generating sketches than in generating composites. The moving around may result in better relationships (e.g., distances) between features than a process oriented toward completing work on one feature prior to moving to another. Of course, the nature of the Identi-kit makes this latter, feature-oriented procedure more likely.

The fact that there was little difference between images from description and view with the Identi-kit has an interesting implication. It may be that a major limiting factor in the quality of composites is the Identi-kit itself, not the ability of technicians. This idea is further supported by the fact that there was little or no difference between technicians, whereas there were artist differences.

The results of the most-attended-to-features analysis (Table 3) seem to reflect a tendency to give more time and attention to features in the upper half of the face than in the lower half. This outcome is consistent with several previous findings (Ellis et al., 1975; Goldstein & Mackenberg, 1966; Laughery, Alexander, & Lane, 1971). It may be, as Goldstein and Mackenberg suggest, that the upper half of the face conveys more information helpful to memory.

The lack of any learning effect with artist/technician experience may be due to asymptotic performance being realized in the preexperimental training. On the other hand, it may be that using slightly over 20 images was not sufficient experience within which to examine improvement. This latter explanation seems unlikely, however, since learning effects in such tasks usually show up early. There is one additional point to be noted on this matter; namely, the best sketches were produced by BM who was the best trained portrait artist. This finding is consistent with an experience effect.

The problem of obtaining a facial image from a person's memory is difficult at best. This research on the sketch artist and Identi-kit indicates that these procedures are considerably short of perfect. The outcome of the study leads to several conclusions that have implications for the use of these procedures in law enforcement and basic memory research.

One obvious conclusion is that when a choice between the sketch artist and Identi-kit is available, law-enforcement agencies are advised to obtain sketches. Indeed, our conclusions about the Identi-kit at this point are similar to those of Ellis et al. (1975, 1978) regarding the Photofit; namely, the kit may be of limited value in actual practice. This conclusion is tempered somewhat by the earlier noted fact that recent changes in the Identi-kit may lead to better representations. Also, a basic purpose of the Identi-kit is to eliminate nonsuspects as well as to suggest potential suspects. In this regard, composites might be useful for eliminating unlikely candidates in mug files or lineups prior to exposing them to witnesses. This application may be important, since previous work by Laughery et al. (1971) and Davies, Shepherd, and Ellis (1979) has shown a decrement in recognition as more distractor faces are seen prior to the target face.

The fact that sketches produced better representations is consistent with the results of the study of Davies, Ellis, and Shepherd (1978). Using a task in which subjects attempted to identify images of well-known faces, three types of images were used: photographs, detailed line drawings, and outline drawings (only the outline of the face and main internal features). Photographs were significantly more effective than line drawings, which in turn were superior to outlines. A conclusion to be drawn, of course, is that face construction systems that employ more lifelike features and/or greater detail are likely to be more useful. Whereas the work on the Photofit indicates that a simple assemblage of photograph parts is not necessarily a productive approach, other systems that blend photographic features, such as the Minolta Photomontage Synthesizer (Duncan & Laughery, Note 2), may prove promising.

Clearly there are factors other than the technique that contribute to image quality. Artist skills matter (technician skills might matter in some circumstances, but the data

here did not show it), and efforts should be made to maximize these skills through selection and training. An interesting and related development in this area is the work of Gillenson and Chandrasekaran (1975). They have developed a computer-based system with which the nonartist witness can create, on a graphic display, male Caucasian facial images. The computer system contains prestored facial features, an average face that serves as a starting point, and a heuristic strategy that guides the witness through a process of modifying the facial image to produce the desired outcome. The potential of such a system is partly in the elimination of the artist/technician and the verbal description from the construction process. Clearly, these two aspects of the sketch artist and Identi-kit techniques are sources of error.

The positive correlations between images from view and from description imply that some faces are more difficult than others to represent with these techniques. The basis of such target effects is unclear at this time.

The correlations between image quality and witnesses' imagery and SAT scores produced one potentially interesting result. The −.487 correlation between the SAT Verbal scale and the composites suggests that people with better verbal abilities may be able to provide better descriptions, which in turn lead to better images. Although this finding does not argue for the use of the SAT in assessing potential witnesses, it does suggest that further research on the verbal description aspect of these techniques may be worthwhile.

A final comment concerns the generally regarded poor quality of images produced with artists and the Identi-kit as well as with the Photofit (Ellis et al., 1975, 1978). Whereas there may be many reasons for the limitations of these techniques, two points deserve further mention. The first point concerns the verbal description dimension of the tasks. Given that people are not good at verbally representing and describing faces (Davies, Shepherd, & Ellis, 1978; Shepherd, Davies, & Ellis, 1979), any procedure that relies heavily on such descriptions will be limited. Possible solutions to this dilemma would be the development of construction techniques that minimize or eliminate the need for description, that is, procedures that permit the witness to construct the face without an artist or technician. Alternatively, procedures might be developed that incorporate standard descriptors or terminology that the witness could effectively adopt. The second point is more theoretical and concerns our understanding of human memory for faces. Baddeley (1979) has argued that facial memory may primarily involve a global or holistic strategy. The present results are certainly consistent with this argument. Since the sketch artist and Identi-kit are essentially oriented toward a feature-by-feature construction, and a global or holistic-processing approach would be less useful than a feature-analysis strategy, the low quality might be expected. Construction techniques that draw more heavily on holistic information (shapes, relationships, etc.) may lead to more highly regarded representations.

Reference Notes

1. Laughery, K. R., Duval, G. C., & Fowler, R. H. *An analysis of procedures for generating facial images* (Report No. UHMUG-2). Houston, Tex.: University of Houston, Psychology Department, 1977.
2. Duncan, F. H., & Laughery, K. R. *The Minolta Montage Synthesizer as a facial image generating device* (Report No. UHMUG-4). Houston, Tex.: University of Houston, Psychology Department, 1977.

References

Baddeley, A. D. Applied cognitive and cognitive applied psychology: The case of face recognition. In L. G. Nilsson (Ed.), *Perspectives on memory research*. Hillsdale, N.J.: Erlbaum, 1979.

Davies, G. M., Ellis, H. D., & Shepherd, J. W. Face recognition accuracy as a function of mode of representation. *Journal of Applied Psychology*, 1978, *63*, 180–187.

Davies, G. M., Shepherd, J. W., & Ellis, H. D. Remembering faces: Acknowledging our limitations. *Journal of the Forensic Science Society*, 1978, *18*, 19–24.

Davies, G. M., Shepherd, J. W., & Ellis, H. D. Effects of interpolated mugshot exposure on accuracy of eyewitness identification. *Journal of Applied Psychology*, 1979, *64*, 232–237.

Ellis, H. D., Davies, G. M., & Shepherd, J. W. A critical examination of the Photofit system for recalling faces. *Ergonomics*, 1978, *21*, 297–307.

Ellis, H. D., Shepherd, J. W., & Davies, G. M. An investigation of the use of the Photofit system for recalling faces. *British Journal of Psychology*, 1975, *66*, 29–37.

Gillenson, M. L., & Chandrasekaran, B. A heuristic strategy for developing human facial images on a CRT. *Pattern Recognition*, 1975, *7*, 187–196.

Goldstein, A. G., & Mackenberg, E. J. Recognition of human faces from isolated facial features: A developmental study. *Psychonomic Science*, 1966, *6*, 149–150.

Laughery, K. R., Alexander, J. F., & Lane, A. B. Recognition of human faces: Effects of target exposure time, target position, pose position, and type of photograph. *Journal of Applied Psychology*, 1971, *51*, 477–483.

Richardson, A. *Mental imagery*. New York: Springer, 1969.

Shepherd, J. W., Davies, G. M., & Ellis, H. D. The relative effectiveness of ratings and verbal descriptions in the recognition of faces. In M. E. Gruneberg & P. E. Morris (Eds.), *Practical aspects of memory*. New York: Academic Press, 1979.

Received September 10, 1979 ■

Manuscripts Accepted for Publication

Validity Generalization Results for Tests Used to Predict Job Proficiency and Training Success in Clerical Occupations. Kenneth Pearlman (Personnel Research and Development Center, U.S. Office of Personnel Management, 1900 E Street N.W., Washington, D.C. 20415), Frank L. Schmidt, and John E. Hunter.

The Unmeasured Variables Problem in Path Analysis. Lawrence R. James (Institute of Behavioral Research, Texas Christian University, Fort Worth, Texas 76129).

The Estimation of the Predictive Power of a Regression Model. Philippe Cattin (Box U-41M, Department of Marketing, University of Connecticut, Storrs, Connecticut 06268).

Person–Situation Effects in the Prediction of Performance: An Investigation of Ability, Self-Esteem, and Reward Contingencies. James R. Terborg (Department of Psychology, University of Houston, Houston, Texas 77004), Peter Richardson, and Robert D. Pritchard.

The Situational Review. Gary P. Latham (Management & Organization DJ-10, University of Washington, Seattle, Washington 98195), Lise M. Saari, Elliott D. Pursell, and Michael A. Campion.

A Preference Mapping of Organizational Objectives of Sports Franchise Executives. David W. Stewart (Owen Graduate School of Management, Vanderbilt University, Nashville, Tennessee 37203) and Donald R. Latham.

Studies on the Perceived Predictive Accuracy of Biorhythms. Robert E. Prytula (Department of Psychology, Middle Tennessee State University, Murfreesboro, Tennessee 37132), Cyril J. Sadowski, Joen Ellisor, Danelle Corritore, Rene Kuhn, and Stephen F. Davis.

Aging and the Development of Automaticity in Visual Search. David J. Madden (Center for the Study of Aging and Human Development, Box 2980, Duke University Medical Center, Durham, North Carolina 27710) and Robert D. Nebes.

Job Choice: The Impact of Intrinsic and Extrinsic Factors on Subsequent Satisfaction and Commitment. Charles A. O'Reilly III (School of Business Administration, 350 Barrows Hall, University of California at Berkeley, Berkeley, California 94720).

Intraorganizational Influence Tactics: Explorations in Getting One's Way. David Kipnis, Stuart M. Schmidt (Department of Industrial Relations and Organizational Behavior, Temple University, Philadelphia, Pennsylvania 19122), and Ian Wilkinson.

Commitment to the Union: Development of a Measure and an Examination of Its Correlates. Michael E. Gordon (Department of Management, University of Tennessee, Knoxville, Tennessee 37916), John W. Philpot, Robert E. Burt, Cynthia A. Thompson, and William E. Spiller.

Intrinsic Motivation as Influenced by Rewards, Task Interest, and Task Structure. Thomas L. Daniel and James K. Esser (Department of Psychology, Lamar University, P.O. Box 10036, Beaumont, Texas 77710).

Exploratory Comparative Study of Four Job Analysis Methods. Edward L. Levine (Department of Psychology, University of South Florida, Tampa, Florida, 33620), Ronald A. Ash, and Nell Bennett.

Individual Correlates of an Occupational Stereotype: A Reexamination of the Stereotype of Accountants. Andrew S. Imada (Assistant Professor, Education Center, USAG—Yongsan, APO, San Francisco 96301), Clive Fletcher, and Anthony Dalessio.

Expectancy Theory Prediction of Goal-Theory Postulate, "The Harder the Goals, the Higher the Performance." Tamao Matsui (Rikkyo University, Nishi-Ikebukuro, Toshima-Ku, Tokyo 171, Japan) and Reiji Mizuguchi.

(Continued on page 345)