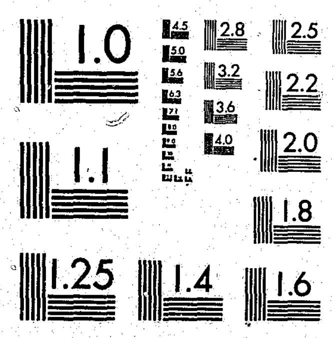


MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

5-27-82

Final Report on Phase One of the Project:

"A Man-Computer System for Solution of the Mug File Problem"

LEAA Grant 74-NI-99-0023 G

FINAL REPORT ON PHASE ONE OF THE PROJECT

"A MAN-COMPUTER SYSTEM FOR SOLUTION OF THE MUG FILE PROBLEM"

Prepared for the Department of Justice, Law Enforcement Assistance Administration, National Institute of Law Enforcement, and Criminal Justice, under Grant 74-NI-99-0023 G

By

Ben T. Rhodes, Jr.
Kenneth R. Laughery
James D. Bargainer
James R. Townes
George W. Batten, Jr.

University of Houston

Table of Contents

Executive Summary 1

Computer Systems for Handling Mug Files

Introduction 5

Objectives

Human Factors 10

Image Generation Study
Method
Results
Discussion
Accessories Study

Hardware 31

Present Hardware Configuration
Hardware System for Police Departments
Phase II

Pattern Recognition 38

Introduction
Design Philosophy
Data Base
The LA Algorithm
Results
Phase II

Software 51

Forgery Applications 53

Further Research, Phase II 55

Objectives for Phase II

Appendices 57

Executive Summary

The objective of this project is to develop a man-computer system for solution of the mug file problem. The mug file problem can be described as follows. A witness to a crime has an image of a suspect in mind, including information about the suspect's face. A law enforcement agency has a large set of photograhps of faces, called mug shots, in its file. The problem is to find an efficient method to use the information the witness has about the suspect's face to help determine whether the suspect is in the mug file.

The approach used in this project involves four steps: (1) obtain an image of the face from the witness using a sketch artist, and Identi-kit, or a similar device, (2) measure certain parameters of the face on the image obtained from the witness and enter these into a computer program, (3) the computer program searches the parameters of the faces in the mug file and determines which mug shots are "look alikes" to the image supplied by the witness, (4) the investigator and witness examine the selected "look alikes" to determine if one or more of those individuals should be considered a suspect.

The research plan for this project called for the use of an interdisciplinary team over a period of 2.0-2.5 years. Phase I of this plan was completed in the period June 10, 1974 to July 31, 1975 and this work is the subject of this report.

The report is broken down into sections which represent the areas of responsibility of the members of the interdiscplinary team. The faculty members on the team and their areas of responsibility are as follows:

Name	Department	Area
Laughery	Psychology	Human Factors
Bargainer	Electrical Engineering	Hardware
Towns	Electrical Engineering	Pattern Recognition
Batten	Electrical Engineering	Software
Rhodes	Industrial Engineering	Forgery Application

The key findings during Phase I of the project are:

1. The pattern recognition algorithm developed during the project can be used to select the look-alikes from a mug file using the information supplied by either an Identi-kit image or a sketch artist's drawing.
2. Inexpensive mini-computers and hardware can be used to implement this approach in a law-enforcement agency.
3. Changes in a person's appearance would not hinder the computer program in selecting look-alikes; however, they can be expected to have a negative effect on the ability of the witness to recognize the suspect.

The relation of the work in this project to other work includes the following:

Computer systems for handling mug files

Several law-enforcement agencies have devices for sorting a mug file on descriptors such as age, height, weight, type of crime, etc. Some of these, such as the Miraquic system developed at Queens, New York City P.D. have proved to be effective

tools for working with a witness. Another system, which was completed about the time this project began, is the CRIME system developed by the Oakland, California P.D. This system uses a modern mini-computer and microfiche display device. It provides for sorting the mug file on a variety of descriptors. Systems like these permit the computer to construct an instant "mug book" but they do not use information about the face such as is available in Identi-kit images and artist's sketches.

In the pattern recognition area, Bledsoe and Hart were early workers in the U.S.A. as far as application to human faces. Kaya and Kobayashi of Japan and Harmon in the U.S.A. have published basic papers recently.

In the human factors area Laughery and others have published the results of a varity of experiments on the recognition task of the witness. The Identi-kit training materials are concerned with interviewing the witness to generate an image. More recently, researchers have designed interactive computer systems to generate images of faces.

Implementation of the results of Phase I of this study may require more development than most law-enforcement agencies are prepared to do. Those which have computer systems could obtain our pattern recognition program and begin to use it. If they have a descriptor sorting procedure similar to Oakland's CRIME system, they would probably wish to combine these two tools. We plan to do this (combine the capabilities of our program and the Oakland program) in Phase II.

Other problems which need further research to increase the potential usefulness of this approach include an efficient procedure to "measure" a mug shot and enter this data into the computer. Better training procedures for sketch artists and better devices to obtain a facial image from a witness would provide better information to begin the search. There are many possible procedures which can be used to display the "look alikes" to the witness, and we need to know which procedures improve the probability of recognition. The forgery application requires development of techniques to "measure" facial images which are taken at different angles.

Since this report covers only Phase I of a project, it will be of primary interest to others who are doing research in these areas and to individuals in the law-enforcement community who make it their business to keep up with new developments. The appendix contains drafts of a paper which is being submitted for publication by members of the project team. This will help disseminate our findings to other researchers. Most of the reports appropriate for dissemination to law-enforcement agencies will be prepared during Phase II.

Introduction

Most of the objectives for Phase I outlined in the original proposal have been achieved. In addition to the original objectives a variety of additional objectives have been identified; some have been achieved and others are being studied. A brief discussion of the status of each objective is given below.

Objectives

1. Selection of the best method of obtaining the image the witness has in his memory in a form which can be used by the pattern recognition algorithm of the system. This will include evaluation of at least three approaches: the sketch artist; the Identi-kit; the Minolta Photo Montage Synthesizer. (Phase I, Human Factors)

This objective has been achieved. Our data indicates that the images obtained from a sketch artist are preferred to Identi-kit images; however, both techniques provide satisfactory images for use in our system. The Minolta Photo Montage Synthesizer is not a competitive procedure at this point in its development. We are currently making hardware modifications and trying to develop a "software system" which will make this procedure competitive.

2. Quantitative measurement of the influence on the probability of correct identification caused by differences in the suspect's accessories at the time the mug shot was made

and the time of the crime. This will include controlled studies of factors such as glasses, mustaches, hair styles, etc. (Phase I, Human Factors)

Laughery and Fowler have completed an experiment which measures these effects for glasses, beards, and long versus short hair styles. They found a marked negative effect upon recognition, with hit rates dropping as much as 42 percent. A draft of the paper on this study that they have prepared for publication is included in the appendices.

3. Development of a computer algorithm which will compare the image supplied by the witness to the photographs in the mug file and select all the "look alikes". The algorithm will by capable of ordering the "look alikes" according to similarity to the image supplied. An effort will be made to make the algorithm insensitive to the age of the photograph in the mug file. (Phase I, Pattern Recognition)

This algorithm has been completed. A brief description of it is given in the Pattern Recognition section of this report. A draft of the paper on this work that Dr. Townes has been prepared for publication.

4. Adaptation of the computer and laboratory facilities at the University of Houston necessary to accomplish the research on the mug file problem. This will include methods of getting various kinds of images into our computer and displaying output images. (Phase I, Hardware)

This objective has been achieved. A description of the facilities in the Image Analysis laboratory at the University of Houston is given in the Hardware section of this report.

This completes the Phase I objectives in the original proposal. The remaining original objectives were scheduled for Phase II. A more accurate description of the current objectives of Phase II is given in the Phase II proposal; however, a brief discussion of the original objectives is included here for completeness.

---

5. Quantitative measurement of the influence of physiological changes such as normal aging and typical changes in weight on the witness' ability to recognize an out-of-date photograph. (Phase II, Human Factors)

The priority of this objective is lower than some of the new objectives which have been established, but it is still in the plan as Milestone 17 of Phase II.

6. Development of methods to "up-date" a mug file photograph and produce a simulated mug file photograph with the accessories specified by the witness. (Phase II)

There are two possible approaches to this problem being considered. One invloves use of the Minolta Photo Montage Synthesizer. While this device is not to the point it can be used for the initial image generation task with the witness, it is being used to modify photographs for some accessories. The other approach is to modify a digitized image in the memory of the computer. One graduate student is currently working in this area.

7. Development of methods to simulate changes to the face image which occur due to physiological changes such as normal aging and weight change. The relative importance of this step and the precision required will be better defined by the studies leading to Objective 4. (Phase II)

The same approaches applied to Objective 6 above, can be applied to this objective; however, we have made this objective a very low priority for two reasons. The main reason is that our computer algorithm has been designed so that these factors should not influence its performance. The second reason is that most experienced law enforcement people tell us this is not important in field applications. When we get the data from Milestone 17 of Phase II, we may reconsider this question.

8. Development of an inexpensive, special purpose mini-computer and peripheral equipment which can be used to sort a large mug file and display images to a witness. (Phase II, Hardware)

This objective has been achieved by combining mini-computers and peripheral equipment currently available. A description of the Oakland, California system and the current cost for it are included in the Hardware section of this report. We will demonstrate a similar system in Phase II.

9. Development of an efficient and inexpensive method of converting existing mug files in law enforcement agencies into

a format which can be used by the man-computer systems such as the one in Objective 8 above. This may also include secondary objectives such as the economies of space with microfilm storage. (Phase II, Pattern Recognition)

This is a major objective for Phase II. We are building a "light pen" device which we expect to be one approach to this problem.

10. Evaluation of the entire system and procedure as a tool for law-enforcement agencies. (Phase II)

This is a major objective of Phase II. Many activities will be directed at evaluation of our system, but the true test must come fron applications by law-enforcement agencies. We hope to begin a field demonstration before the end of Phase II.

Human Factors

A central issue to the mug-file problem in criminal identification concerns the memory that the witness has of the target person. The typical use of mug files actually involves the witness' memory at two stages of the process. The first memory task occurs when the witness initially encounters the identification system. This task involves an effort to recall some characteristics of the target in order to reduce the size of the file. For example, the witness may note that the target was a white male, thus permitting black males and all females to be eliminated from the set of alternatives. The second stage involving memory is the recognition task, where the witness is looking at pictures of faces and making decisions about whether or not each face is the target person.

In the first phase of the project the human factors activities focused upon two studies. The first study dealt with the initial part of the identification task; namely, the attempt by the witness to recall characteristics of the target. This study, referred to as the image generation study explored procedures for generating visual images of the target. Two techniques were examined; sketch artists and the Identi-kit.

The second study was related to the recognition task. An experiment was carried out to explore the effects of

changing accessories between initial exposure to the target and the target's appearance in a subsequent recognition task. Specific accessories manipulated in the experiment were glasses, beards, and hair styles.

<u>Image Generation Study</u>

Law enforcement techniques in the past have included several image generation procedures. Sketch artists and the Identi-kit are two of the most common techniques in use today. The sketch artist technique, as the term implies, 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 facial feature. There are a large number of sheets for each feature; e.g. many types of noses, eyes, etc. A trained technician constructs a composite face by interacting with a witness to select appropriate features. The present study explored the sketch artist and Identi-kit procedures as means of obtaining a target image from a witness.

Actually, there are several purposes or goals of the work reported here, and they are reflected in the following questions:

1. How accurate an image can be generated with sketch artist and Identi-kit procedures? What do the distributions of this accuracy look like?
2. What are the relative merits of the sketch artist and Identi-kit techniques?

3. How much effect does the artist or technician have on the accuracy of an image?
4. What characteristics of the witness influence image accuracy and to what extent?

In addition to seeking answers to the above questions, our efforts were directed towards developing improved techniques and procedures for using sketch artists and the Identi-kit. Limited changes or modifications in procedure were introduced as the study progressed. Most of these modifications were related to the nature of the interaction between the artist or technician and the witness.

Method

As already noted, this study is intended to address a number of questions and issues. The design and procedures are not straightforward. In part, the design consists of manipulating several controlled variables in a manner that falls neatly into an analysis of variance research model. Measures on other variables were obtained, however, with the idea of correlating them with various performance or outcome measures. In this section we shall first describe the basic design of the image generation part of the experiment and then note the other measures that were obtained.

The subjects in this study can be divided into two groups, those who served as targets and those who served as witnesses. A total of 97 target subjects were used, all white males. The

targets were drawn from several sources, including students at the Universtiy of Houston and the Houston commuity at large. The only restriction placed upon the selection of these subjects was that they be unknown to the witness subjects, the sketch artists and Identi-kit technicians. These were 182 witness subjects. All subjects were paid $2.00 per hour for participating.

There were two phases in the basic image generation task. The first phase consisted of a conversational interaction between the witness and target. This interaction followed instructions to the witness that he/she would subsequently be working with a sketch artist or Identi-kit technician to create the target image.

Two variables were manipulated in the design of the study The first was the image-generation technique, consisting of the sketch artist and the Identi-kit. The second variable will be referred to as artist-technician, which represents three artists and three Identi-kit technicians. The artist-technician variable was nested within technique; that is, the three artist and three Identi-kit technicians were six different people. Because the training and ability of these six people is crucial to the study, a brief summary of their credentials is presented in appendix HF1.

As stated above, 182 witness subjects and 97 target subjects were used. The manner in which witnesses and targets were paired and the assignment of witnesses to artists and technicians was not balanced. It should be noted that in terms of "purity", certain types of confounding is unavoidable given that a particular target cannot be described more than once by a witness nor constructed more than once by a given artist or technician. The actual pairing of targets and witnesses and the assignment of witnesses to artist-technicians was done in the following manner. An effort was made to have each target exposed to two witnesses, one of whom then described him to an artist and the other to a technician. We were successful in this regard for 78 targets; that is, there were 78 targets each exposed to two witnesses, each done by both an artist and a technician. It was not possible to balance the artists and technicians with respect to targets. The following Table shows the number of targets shared by the different combinations of artists and technicians.

Number of Targets Completed by Different Combinations of Artists and Technicians

		Sketch Artist			
		RM	SN	AM	Total
Identi-kit Technician	MM	15	4	5	24
	RF	5	14	9	28
	JH	4	6	16	26
Total		24	24	30	78

The remaining 19 targets and 30 witnesses were paired and assigned to insure that each artist and technician constructed a minimum of 30 images. In several cases, two witnesses worked on the same target, but also the same technique. The number of completed sketches was 92 and Identi-kit composites was 90.

The procedural aspects of each regular experimental session of the study involved six people: the experimenter (E), a sketch artist (SA), and Identi-kit technician (IKT), a target subject (TS), and two witness subjects (WS). Since it was necessary to carefully control the timing and manner in which different individuals encounter each other, and because a variety of data was collected from the various individuals, a relatively complex and carefully controlled procedure was carried out. The specific steps were as follows:

1. Two witness subjects reported to a room where they were met by E. Upon their arrival they were asked to complete a Subject Data Form which required a total of approximately five minutes. This form asked for information about the S, including certain physical characteristics. A copy of the form is presented as exhibit 1 in appendix HF2.
2. After the data forms were completed, photographs were taken of each WS. The photographs included front, left profile and right profile bust-length

views. If the WS wore glasses, two front views were taken, one with and one without the glasses.

The photographs were taken with a half-frame Olympus 135 mm. camera with Ektcrome film. Acutally the film was made into slides, not prints. For purposed of this report, however, samples of the pictures made for a WS have been printed and are presented as exhibit 1 in appendix HF3. The physical parameters of all slides were constant (sharpness, scale, lighting, etc.).

3. After the photographs were taken, the two WSs were instructed by E as to the nature of the experiment. A sample set of instructions shown in exhibit 1 of appendix HF4. This is a sample in the sense that E did not read the instructions; they were presented in a conversational fashion (having been well rehearsed).

4. While the two WSs were completing the data forms and being photographed, the TS reported to an adjacent room. After E finished with the WSs, he greeted the TS and presented instructions regarding the study. These instructions are shown as exhibit 2 in appendix HF4 and were also delivered in a conversational manner.

5. Following the instructions, E escorted the WSs to the room where TS was waiting. It should be noted that all three Ss at this point were aware of the nature of the experiment and the nature of the image generation task. The E, TS and WSs were seated at a table (TS across from the WSs). The E then moderated a 7 to 8 minute conversation among the subjects, hereafter referred to as the exposure period. To the extent possible, the discussion focused upon TS: what was his major (if student) or job; where did he live; what were his interests; etc. A sample of Es introductory remarks in this session is presented as exhibit 3 in appendix HF4. While the setting may seem somewhat strained or artificial, in actual practice it generally proceeded quite smoothly with reasonably good conversation.

6. After the exposure period, one WS was escorted to a room to work with a sketch artist to generate an image, while the second WS was taken to a room to work with and Idneti-kit technician. Upon arriving in these rooms, the WSs initially filled out a general discription form about the TS. This form called for information about TS that was used by the sketch artist or technician as a starting point for generating the image. The forms used in the two techniques were slightly different, and are shown as exhibits 2 and 3 in appendix HF2 for the sketch and

Identi-kit techniques respectively.

After completing the general information forms, the WSs worked with the artist/technician to produce the image. The verbal interaction in each situation was tape recorded using a Stenorette Embassy dictating machine. A sample of the sketch from description, sketch from view, composite from description and composite from view are included as exhibits 2,3,4 and 5 respectively in appendix HF3.

7. While the WSs were working on the image generation task, TS completed the Subject Data Form, exhibit 1 in appendix HF2.

8. After completing the Subject Data Form, TS posed for photographs. The same pictures were taken of TS as described above for the WSs.

9. After the WSs finished the image generation task, they completed three additional forms. The first was a Subject Comments Sheet. This form solicited comments form WSs regarding the manner in which they carried out the task. The form is presented as exhibit 4 in appendix HF2.

The second and third forms consisted of the Betts and Gordon tests for imagery ability. Both are paper and pencil procedures for assessing ability to carry out imagery or verbal memory activities.

Samples of the Betts and Gordon are presented as exhibits 1 and 2 in appendix HF5, respectively.

10. While the WSs were completing the three forms described above, TS reported to a room where the sketch artist and Identi-kit technician produced a sketch and composite of TS while viewing him directly.

Results

The following list summarizes the data collected in this study:

1. Photographs of TS and WS.
2. Sketch of TS from WS description.
3. Sketch of TS from direct artist viewing.
4. Identi-kit composite of TS from WS description.
5. Identi-kit composite of TS from direct viewing.
6. Recorded protocols of the verbal interaction between WS and artist or technician.
7. Information on TS and WS contained in Subject Data Form.
8. Scores on Betts and Gordon Imagery tests.
9. WS answers to questions on Subject Comment Sheet.
10. Answers to questions on Interview Procedure Form.
11. SAT verbal and quantitative scores on subjects who were undergraduate students at the University of Houston.

The data analyses of the image generation study are not yet complete. Several preliminary analyses have been carried out, however, and these results are presented in this section.

An important and nontrivial set of issues in this study concerns the manner in which one compares facial images. More precisely, what are the dependent measures or criteria by which performance is evaluated? To date, our efforts have proceeded in several directions. One of these directions involves measures of the physical dimensions of the face, while the second employed subjects' ratings of similarity.

The various physical measures of the face are described in the Pattern Recognition part of this report on page 82. The different facial images generated in the study were not carried out to the same scale. That is, the photographs, sketches and Identi-kit composites produce images of different sizes. Indeed, a variety of image sizes were produced just within the sketches. For this reason it is not possible to make comparisions across image types on the basis of the linear dimensions. Instead, a number of ratios of the different dimensions were employed as dependent measures. The specific ratios used in the analyses to date are presented in the Pattern Recognition part of the report on page 82.

A first and rather straightforward look at this data involved a variance analysis of the technique (sketch vs Identi-kit) and different ratio effects. The dependent measure was a difference score: the difference between the ratio value for the photograph and the sketch or Identi-kit. The sketches and Identi-kit composites used here were, of course, those

generated from the WS description. The means for the 32 ratio-technique conditions are presented in Table 3. The analysis of variance showed the main effects of techniques and ratio and the technique x ratio interaction were all significant at the $p < .001$ level.

The data underlying these interactions indicate that performance was better on the sketches; that is, the difference between the sketches and photographs with respect to the ratios were smaller than the Identi-kit photograph differences. The data do not provide a clear interpretation of the ratio or ratio x technique interaction effects. Obviously, there are accuracy differences for the different ratios, but no pattern emerges to clearly indicate more or less accuracy on various parts of the face.

In addition to the above analysis of variance, we have carried out several additional analyses correlating the ratio difference measures to various other performance measures and characteristics of the WSs. Specifically, these variables have included performance on the imagery tests and SAT scores. The results of these correlations have indicated very little, if any, relationship between the various measures.

The second direction in which our analyses of the image generation study results has proceeded involved the use of similarity ratings. These "psychological measures" of goodness-of-fit involve showing subjects photograph-sketch (or composite) pairs and having them rate the similarity of each

Table 1

Mean Differences Between Ratio Measures on Photographs and Sketches or Identi-Kit Composites

	Technique	
Ratio	Sketch	Identi-Kit
1	.034	.013
2	.011	.057
3	.083	.019
4	-.032	-.050
5	-.049	-.075
6	.068	.054
7	.005	.003
8	-.042	-.039
9	.013	.020
10	-.035	-.042
11	-.055	-.063
12	.005	.002
13	-.040	-.053
14	.117	.137
15	-.022	-.002
16	-.022	-.002

pair on some sort of scale. In actuality, a formal experiment was carried out to obtain these similarity ratings. The experiment is described in the following paragraphs.

The Ss were 30 undergraduate students enrolled in introductory psychology courses at the University of Houston. Class credit was given for participation in the study.

The Ss viewed pairs of slides consisting of the front bust of the target and one of the four images produced by:

1. Sketch artist
2. Identi-kit technician, from description
3. sketch artist's sketch from view
4. Identi-kit from view.

All four images were paired with each of the target slides. The Ss' task was to rate each pair on a six point scale with respect as to how well the image matched the target. The slides were projected so as to be approximately life size on the screen.

The design of the experiment was a 2x2x3 with all factors manipulated as within-S variables. The conditions of the first variable, image generation technique, were sketch artist and Identi-kit. The second variable was image type and refers to the two images produced for each target; from the WS description or while viewing the target. The third variable was artist-technician. As stated earlier, this factor was nested within technique; the three sketch artists and three Identi-kit technicians were different individuals.

Ratings were obtained on a total of 204 pairs from each S. There were a total of 51 different targets whose photograph appeared in four pairs; once each with the sketch from description, sketch from view, composite from description and composite from view. Each of the artists or technicians contributed 34 images, 17 done from description and 17 from view.

For each of the 51 targets, black and white slides of the four images were prepared. Four series of images were constructed in which one image for each target was present. Within the series 1/4 of the slides were of each image type. These were randomized with the constraint that no more than three images of the same type might occur successively.

The apparatus consisted of two Kodak Carousel AV 9000 projectors with 4 to 6 in., F3.5 Zoom Ektamar Lens and Da-Lite projection screen.

The procedure involved running each S individually with order of image series counterbalanced among S's. The screen was located at the front center of the room 8 ft. from the S at a height slightly above eye level when seated. The projectors were at the rear of the room on both sides of the S. The room was darkened to insure good vision of the slides, but with sufficient light to read and mark the answer sheets.

The front bust of the target was projected on the left side of the screen and the image on the right. The sequence of targets remained the same for all S's and each S randomly

began his rating at one of 5 points in the series.

S's were instructed to use the left end of the scale for the best matches between target and image and the right end for the poorest matches. The S's were instructed that the intermediate points were to be used in rating those pairs which were neither the best nor the worst matches--keeping in mind the meaning of the end points.

During the rating sequence, each slide pair was projected on the screen for approximately seven seconds, and the S marked his reponse on the answer during the two seconds required to reset the projectors. Ss rated 10 pairs prior to beginning the actual ratings used in the analyses. The 10 sample pairs were selected so as to be representative of the range of similarity and included at least 2 of each of the 4 image types.

An anlysis of variance was applied to the results. It should be noted that the technician x method interaction cannot be examined because technician is nested within the method. Table 2 shows the analysis results. All of the effects are statistically significant at a $p<.01$ level.

The data for the various conditions are shown in Tables 3 and 4. Keeping in mind that smaller numbers represent better performance, it can be seen in Table 3 that sketches were judged to be better images than composites. Also, the view condition led to better images than the description, although this diffenence was quite small with the Identi-kit technique.

The technician x image type interaction can be clearly seen in the data in Table 4. The three Identi-kit technicians differ very little form each other, while there is considerable variation among the artists.

Discussion

Although only preliminary analyses have been carried out on the image generation study to date, the results are consistent in showing that sketches are better representations that Identi-kit composites. A great deal of data analysis remains. however, including the following activities currently in progress:

1. A composite measure based upon physical dimensions has been developed and will be used in comparing image techniques as well as in many other comparisons and correlations.
2. Correlations are being computed between the similarity rating values and a number of witness characteristics.
3. Transcriptions of the protocols have been completed and are being analyzed to determine the adjectives used in describing various facial features and the sequence in which different features are dealt with in generating images.
4. Strategies reported by witnesses are being categorized and analyzed with respect to the quality of the images.

Table 2

Analysis of Variance Results: Similarity Ratings in Image Generation Study

Source	df	MS	F	p
Method	1/23	24.01/1.08	22.19	.01
Technician	4/92	3.39/.16	21.33	.01
Image Type	1/23	9.62/.39	24.93	.01
Method x Image Type	1/23	5.03/.30	16.87	.01
Technician x Image Type	4/92	.575/.13	4.56	.01

Table 3

Means of Method and Image Type Conditions from Similarity Ratings in Image Generation Study

		Method		
		Sketch	Identi-Kit	Mean
Image	View	2.92	3.76	3.34
Type	Description	3.55	3.86	3.71
	Mean	3.23	3.81	

Table 4

Means of Technician and Image Type
Conditions from Similarity Ratings
in Image Generation Study.

		Technician						
		Sketch			Identi-Kit			
		1	2	3	4	5	6	Total
Image	View	2.56	2.70	3.50	3.67	3.85	3.76	3.34
Type	Description	3.32	3.56	3.74	3.91	3.91	3.76	3.71
	Total	2.94	3.13	3.62	3.79	3.88	3.76	

Accessories Study

A study has been completed which examined the effects of changing accessories between WS's initial exposure to the target and the target's appearance in a subsequent recognition task. The accessories studied were glasses, beards, and hair style.

Since a paper dealing with this study was prepared and submitted for publication in the Human Factors Journal, a description of the work will not be presented here. Instead, a copy of the paper is included as appendix HF6.

Hardware

The charge given the hardware group during Phase I was to develop the necessary hardware to support the other phases of the research and to specify hardware which would be cost effective for installation in a police department. This has been done and certain additional problems have been defined and solved.

<u>Present Hardware Configuration</u>

The Image Analysis laboratory computer is configured as shown in Figure 1. The primary computer is a Hewlett Packard 2100A mini computer with 16K, 16 bit words of core memory. It is interfaced to a 1/2", IBM compatable, 9-track magnetic tape and disk with a fixed platter and a removable one. The total capacity is 5 million bytes (8 bits/byte). A Mini Bee II CRT terminal is used to interact with the computer and a teletype is used to obtain hard copy output. The other Program I/O device is a paper tape punch and reader which is shared with the SDS92 Computer. In addition, a Bensen-Lehner stepper motor type X-Y plotter has been interfaced to the HP2100 to allow the automatic plotting of images and graphs. This computer forms the basic component for the sorting of mug shots and will be interfaced to a random access microfiche display unit during the early part of Phase II.

Of this configuration, the plotter, magnetic tape and CRT are useful for the research phase of this grant, but

would not be included in a system which would be installed in a police department.

To aid in the research, another computer, the SDS92, has been interfaced to the HP2100. This machine is used primarily as an off-line processor for digitizing images from a closed circuit television system. The SDS92 is interfaced to the TV system for digitizing the images on a (128 x 128) or (256 x 256) array. It can display the digitized image on a Teletronix 611 storage display. Final construction is being done on electronics which will allow the display of the digitized images on the television monitor.

Software has been written for controlling the SDS92 when digitizing images. This software has been called IMAGE and offers a special purpose language with the following instructions:

SRLO	Set register to low resolution (128 x 128)
SRHI	Set register to high resolution (256 x 256)
SCAN	Digitize the image on the monitor at the existing resolution
DISP	Display the image on the storage scope
HOLD	Place the storage scope in the HOLD mode
VIEW	Place the storage scope in the VIEW mode
ERSE	Erase the storage scope

In addition, there is an executive language which allows the printing of a core dump, changing of core, etc.

The SDS92 is interfaced to the HP2100A so that the data can be transferred to or from the HP2100A and the HP2100A can force execution of predetermined routines in the SDS92. An image can be digitized by the SDS92 and this image transferred to the HP2100. The HP machine can then perform operations on the image and transfer the image back to the SDS for display on the stroage scope. This is very useful for analyzing images and operations such as converting images to line drawings or developing algorithms to transform images to normalized positions, or developing algorithms for automatic measurement of parameters on the image.

When a police department wants to install a system like ours they will need to "measure" a large number of photographs. To facilitate this, we have developed a light pen/joy stick system. Cross-hairs are positioned by light pen or joy stick or a combination of the two. The X,Y coordinates of the intersection of these cross-hairs are then entered automatically into the computer upon command. This should greatly increase the speed and accuracy of taking the measurements. This system is in the final construction phase.

The hardware design projects have been done by students superyised by Dr. Bargainer and are shown below.

A. Design of IMAGE Hardware and Software consisting of the hardware and software required to digitize the image and display it on the storage scope was done by Mr. Gary Hornbuckle, a graduate student

majoring in Electrical Engineering and supported by the Electrical Engineering Department.

B. Design of the plotter interface was done by Mr. Mark Franklin, an undergraduate student in Electrical Engineering. Mark was supported on this grant.

C. Design of the SDS92-HP2100 hardware and software was done by Mr. Bernard Gordan, a student in the fifth year design sequence (Master Electrical Engineering). Bernie was not supported during that time.

D. Design of the TV display system is being done by Mr. Ronald Dockal, a graduate student in Electrical Engineering supported by this grant.

E. Design of the light pen/joy stick system is being done by Mr. Martin Daniels, who is a graduate student in Electrical Engineering supported by this grant.

The Hardware System for Police Departments

The system which would be necessary for performing sorting of mug shots, fingerprints, and vehicle information and random access to other records is shown below. Obviously, the Hewlett Packard minicomputer is not the only computer which can be used. There are many fine systems which could be used, some of which might be less expensive. The installation here at the University of Houston and the one at Oakland, California P.D. are HP2100 systems.

HP2108 CPU with 24K random access memory	$ 7300.00
Two Dual Disks with controllers and 2 spare disc packs 5M words total	25150.00
Periphreal Equipment Time Base Generators Teletypes & Interfaces Paper Tape Reader Cabinet	5360.00
Fiche Display Unit with interface and hard copy printer	9995.00
	$47805.00

Other costs would include the following:

Cost of Fiche	$.10/image
Software	Supplied from this grant
Expendable Supplies Paper tape TTY Paper Access forms Printer paper	
Training of Employees	Done by computer manufacturers and LEAA by separate grant to some group

The software developed on this grant would be available and would be significantly decrease the cost of the system.

The CRIME System in Oakland P.D. cost about $100,000.00, including hardware and software.

Phase II

There are two hardware projects which will be completed during Phase II.

1. Install the microfiche display system and modify if necessary to perform for the final system.
2. Investigate available hardware for automatically digitizing the measurements from the fiche display.

The light pen system developed here is a satisfactory system for our use, but it would be too expensive to install with each system. Other techniques more compatable with the fiche display unit must be investigated.

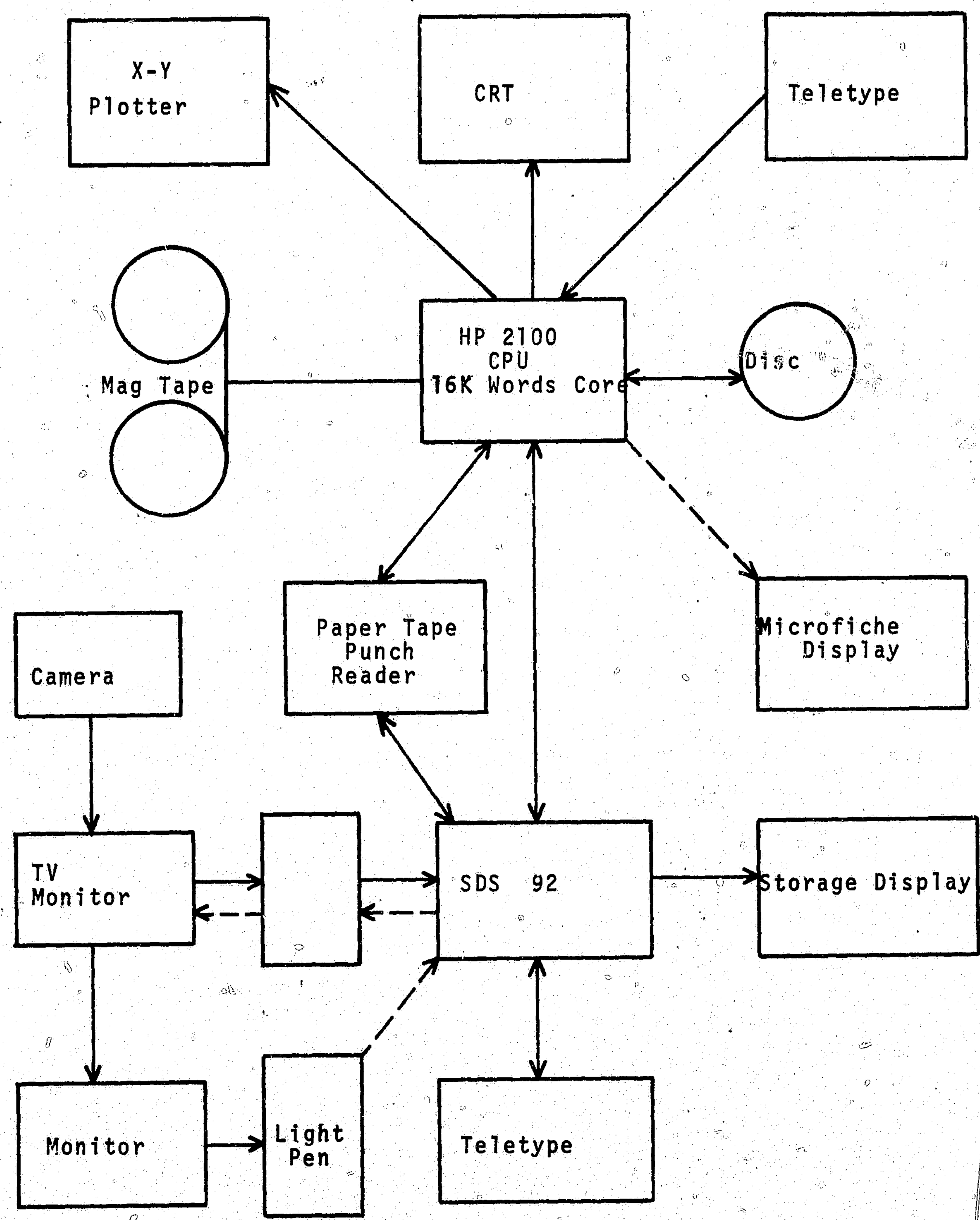


Figure 1

Hardware System

Pattern Recognition

I. Introduction

This section summarizes the philosophy, goals, and achievements during the first phase for the pattern recognition group. The precise details and theorems will be presented in a paper entitled, "A Computer Interface in Mugshot Retrieval" which is being submitted for publication in the IEEE Transactions on Circuits and Systems.

As stated in the original proposal, a major objective was "the development of a computer algorithm which will compare the image supplied by the witness to photographs in the mugfile and select 'look-alikes' (objective 3). Specifically, research during phase I was concentrated on recognition similarities in the basic "shape" of the faces. This point is emphasized in that no qualitative features (glasses, beard, etc.) were used in the look-alikes (LA) algorithm. (During phase II, the LA algorithm will be incorporated with an optional sorting package to restrict the look-alike part of the search to a subset of the mugfile which has already been sorted on qualitative features such as sex, age, type of crime, etc.

II. Design Philosophy

The LA algorithm is designed to permit a witness and a law enforcement agency to by-pass the tedious manual sorting through a mugfile or mug book and to concentrate instead on

a few images which the computer has identified as "look-alikes" to the description given by the witness. How many crimes this will help to solve depends, of course, on the quality of the mugfile and on the ability of the witness to recall and describe the suspect. Thus before any hope of such a real-world system can be realized, attention must be brought to two very basic and distinct questions: 1) On the average, does the witness have the ability to remember and effectively describe a suspect? 2) On the average, can information be extracted from a witness' description and used by a computer to locate the desired photograph?

The results of the research of phase I indicate that both questions can be answered in the affirmative. Data from a simulated mugfile of 100 photographs and 75 sketches is presented to support this claim.

The design of the LA algorithm was accomplished under certain guidelines, or requirements, for its operation. These requirements have been grouped into five basic categories as follows:

R1. First, a measure of similarity between the geometry of one type of facial image must be established. The similarity measure must be able to compare photographs with photographs, sketches with photographs, sketches with sketches, and likewise Identi-kit composites.

R2. The data entered into the computer from a witness' description must be easily measured. It is likely that very few police departments would be willing to invest time in training personnel in comlicated curve tracing or texture measuring. However, it is reasonable to assume that the measuring of distances between certain facial landmarks on a sketch would not be considered too great a task.

R3. Since the algorithm must compare data from a sketch with that of several thousand photographs, in the interest of time (to the user and to the CPU) numerical computations and memory requirements must not be excessive.

R4. The algorithm must not be sensitive to the size of the sketch or to minor variations in the location and orientation of the face in the image.

R5. Finally, and perhaps most important, the algorithm must be effective in matching a witness' description to the correct image. Ideally, the algorithm should consistently select as its first choice the correct photograph or image from the mugfile. Since this cannot be accomplished all the time, objectives can be stated in terms of a computer reduction of the original mugfile population to a small percentage of the images (such as 1%, 5%, or 10%) with a specified high probability of the desired photograph appearing in the reduced set.

III. Data Base

The data which this report is based upon consists of hand-measured features from photographs, sketches, and com-

posites of white, college-aged males. A precise description of this set of images is described in the human factors portion of the report. However, a very brief description will be given here.

From a previous study in Buffalo, New York, a collection of 200 facial images (photographs only) was obtained.

During the phase I study, images from 75 persons were obtained under the following conditions: a target was observed simultaneously by two witnesses (no restriction on age, sex, or the race of the witnesses). After a brief viewing period, one witness interacted with a sketch artist, the other with an Identi-kit operator. After the witnesses were finished, the target appeared before the sketch artist and the Identi-kit operator and another set of images was obtained. Thus, from this study, five images were obtained per target: a photograph (PH), a sketch from witness description (SW), a portrait sketch (SP), and Identi-kit composite from a witness description (IW), and a portrait Identi-kit composite (IP).

A pattern recognition technique called "training" was conducted using images from the 75 targets mentioned above and 25 photographs from the earlier study, thus 100 photographs and approximately 75 sketches of each type were used (not all targets had all five images due to certain scheduling problems).

The numerical characterization of facial images was done in a quantitative, rather than qualitative, manner.

To avoid problems of changing hairstyles, etc., it was decided to restrict all measurements to the area bounded above by the eyes and bounded below by the chin.

As indicated on the drawing, Figure 2, the following measurements were recorded from each image:

M_1= internal biocular distance

M_2= external biocular distance

M_3= nose width

M_4= mouth width

M_5= distance across the face measured directly under the nose

M_6= distance across the face measured across the mouth

M_7= nose length from tip of nose to midline of the eyes

M_8= distance from chin to eyes

M_9= distance from lower lip to eyes

In the case of heavy sideburns or other facial hair, certain features (such as M_5 or M_6) are difficult to measure accurately. In such cases, default antomical landmarks were used. In the case of sideburns and beards, a landmark or 1/3 the distance into the sideburn was chosen as a default point for the edge of the face.

To overcome the problem of different images of different sizes, ratios of the above measurements were computed. (While there is a possibility that this procedure might cause two persons whose head sizes were considerably different to appear "similar", the phenomenon was never observed in the data men-

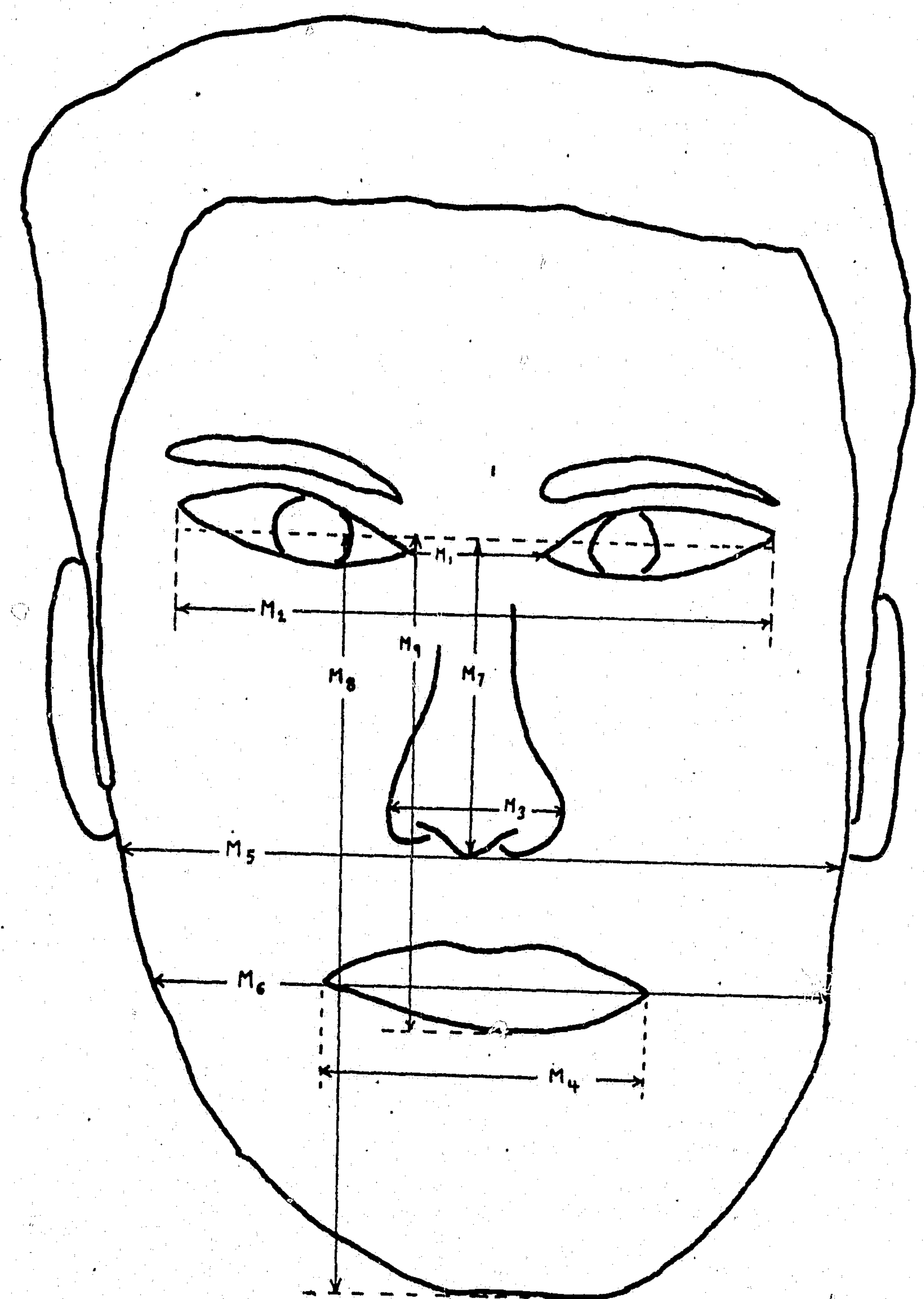


Figure 2

Facial Measurements

tioned above.)

IV The LA Algorithm

In this report, the term "target set" denotes the collection of images which are used as a mugfile. The term "witness set" denotes the collection used (one image at a time) to locate specific image in the target set. The LA algorithm was run on each of the following "target set-witness set" pairs: PH-SW, PH-IW, SP-SW, and IP-IW.

As a preliminary preprocessing step, each ratio from the target set data was normalized to have sample mean equal zero and sample standard deviation equal one.

In addition, witness set ratios were quadratically regressed on the normalized target set ratios. That is, if $W_1, \ldots, W_n$ and $T_1, \ldots, T_n$ represent a specific ratio from n witness images and the normalized ratio from the corresponding target images, respectively, then the constants a_0, a_1 and a_2 are determined by a least-squares fitting of the function $t=a_0+a_1W+a_2W^2$. This is accomplished by selecting a_0, a_1, a_2 to minimize $\sum_i [a_0+a_1W_i+a_2W_i^2-T_i]^2$ where the summation is over the set of target-image pairs (T_i, W_i), i=1, ...,n. The value of W_i is then replaced by $a_0+a_1W_i+a_2W_i^2$.

After each of the 72 ratios chosen from the nine basis measurements are treated in this manner, the eight ratios with the smallest mean square error (the value of the summation given above) were found. These eight are the ratios

which have the closest average agreement between witness image and corresponding target image. (The remaining 64 ratios are not used after this.)

At this point, each image is characterized by 8 ratios $R_1, \ldots, R_8$ (normalized for target images and quadratically regressed for witness images). To bring the eight ratios into closer agreement between target and witness images, a linear processing on witness data is performed:

F = A R

R = collection of 8 regressed ratios

A = matrix of constants

F = collection of 8 processed features where,

$$F_i=a_{i1}R_1+a_{i2}R_2+\ldots+a_{i8}R_8 \quad (i=1,\ldots,8)$$

The processing matrix A can be decomposed into the product of two matrices, $A=A_1A_2$, where:

(1) A is a multilinear regression matrix. That is a collection of constants α_{ij} ($1\le i\le 8$ and $1\le j\le 8$) such that on the average $\alpha_{i1}R_1+\ldots+\alpha_{i8}R_8$ is as close to the corresponding T_i as possible. After these constants are found, R_i is replaced by this sum.

(2) A_2 is a permutation matrix which merely rearranges the processed data in order of each processed ratios' "importance" in classification. This is achieved by the following technique:

F_1=the processed ratio which does the best job of recognizing images.

F_2=the processed ratio which, in combination with F_1, does the best job of facial recognition.

F_i= the processed ratio, which in combination with $F_1, F_2 \ldots$, F_{i-1}, does the best job in facial recognition.

Once the permutation matrix is found, the ratios from both sets of data are recorded.

The concept of data ordering is a most important one. One would expect that, if ordered properly, two features would perform better than one and that three features would out-perform two. However, at some point in this procedure, the amount of additional information gained from a new feature begins to decrease and in the extreme case, adding more features is equivalent to adding "noise". Figure 3 (typical for all target set-witness set pairs) illustrates this phenomenon. The optimal number of features varies according to which data pair was used.

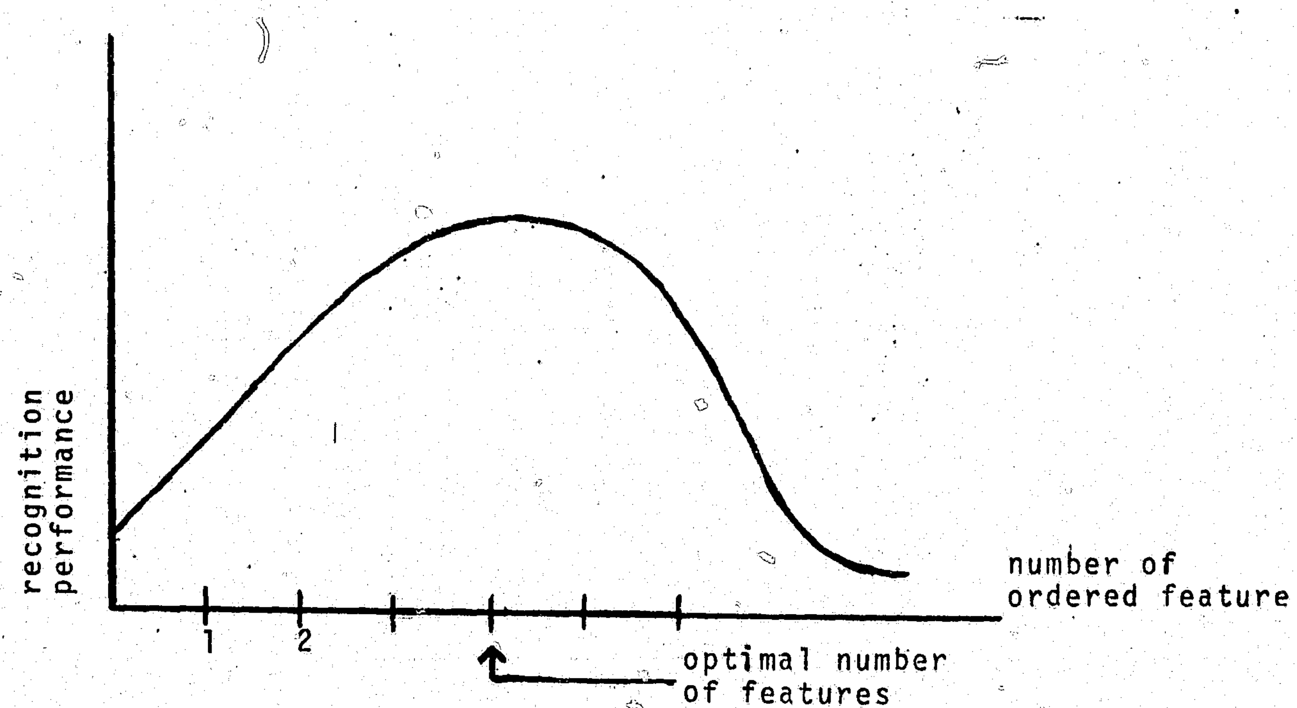


Figure 3

Using the ordered features, a Eculidean distance function is used for evaluating similarity among images. Let F_1, $F_2, \ldots F_n$ and $T_1, \ldots, T_n$ denote the set of processed and ordered features from a witness image and the normalized and ordered ratios from a target image, respectively (N is the optimal number of ordered features). The "distance" between the two images is taken to be $D= [(F_1-T_1)^2 + \ldots + (F_n-T_n)^2]$ Note that a distance of D=0 (which is the smallest possible distance) indicates that there is perfect agreement between the two images. Also as the difference between corresponding values increases, so does the distance between the two images.

To convert a distance D into a similarity rating S, <u>any</u> monotonic decreasing function may be used (that is, if D=0, then S=1 and as D grows large, S goes toward 0). Such a function is S=exp (-D), as shown in Figure 4.

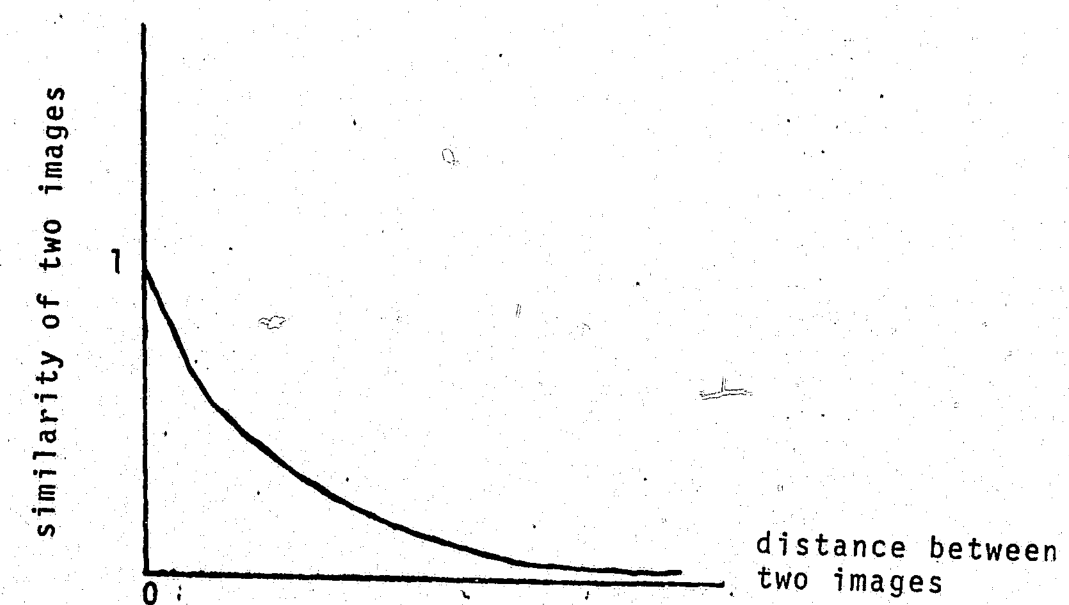


Figure 4

As a ranking procedure, the closest "look-alike" to a given image is the target with the largest similarity rating (or equivalently, with the smallest distance).

V Results

The LA algorithm described above was tested on a data base of 100 photographs (approximately 75 of which had corresponding sketches and composites). The target set-witness set pairs were PH-SW, PH-IW, SP-SW, and IP-IW. The last two pairs were included to measure a possible upper bound in the performance on sketches and composites. (That is, the algorithm should work somewhat better in comparing sketches with sketches than it does in comparing sketches with photographs.)

Table No. 1 lists the eight initial ratios as discussed earlier. For notational simplicity, a ratio of measurement M_i to measurement M_j is denoted by i/j.

Ratio	PH-SW	PH-IW	SP-SW	IP-IW
1	1/5	9/8	7/5	2/6
2	8/2	1/9	1/7	2/5
3	4/7	5/1	5/4	6/4
4	6/1	8/3	7/2	2/8
5	4/9	3/1	7/6	5/6
6	5/8	6/1	1/3	2/3
7	6/4	8/1	5/8	5/4
8	6/7	3/6	1/9	3/6
optimal number of ordered features	4	3	6	7
% of correct first choices	13	10	21	20

Table 1

Figure 5 illustrates the probability of finding the desired image in the reduced target set as a function of the size of the reduced set. The dashed line (performance based on random selection) is included as a reference.

VI Phase II.

The proposed work of the pattern recognition group in Phase II can be briefly summarized as stated below: 1) Combining the present LA algorithm with optional sorting routines 2) extracting or measuring facial features from the T.V. screen using the light pen algorithm (of the hardware section) 3) automatic computer extraction of facial features 4) conversion of photographs into line drawings (this is presently well under way) 5) Testing and improving the overall algorithm on the large mugfile.

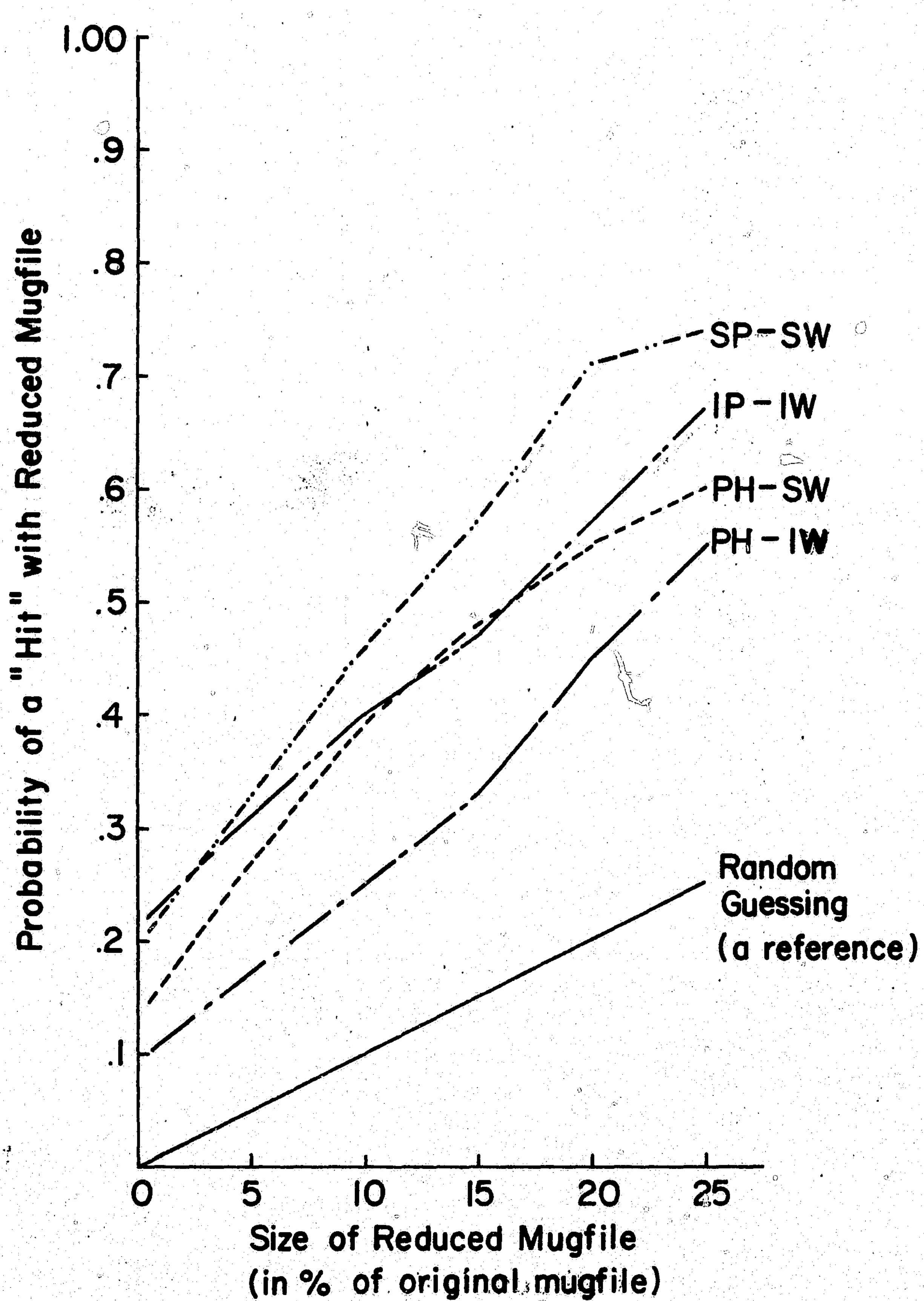

1.00
.9
.8
.7
.6
.5
.4
.3
.2
.1
Probability of a "Hit" with Reduced Mugfile
SP-SW
IP-IW
PH-SW
PH-IW
Random
Guessing
(a reference)
0
5
10
15
20
25
Size of Reduced Mugfile
(in % of original mugfile)

Figure 5

Software

This portion of the Phase I report deals only with software developments other than those relating directly to the development of the "look-alike" algorithm. This work was funded by the LEAA grant only during the last six weeks of Phase I.

Most of the work done during this period was development of supporting computer programs as follows:

1. Modification of HP2100-SDS92 interface drivers and development of interface control subroutines and main program. These have greatly increased the ease of operation of the whole system.
2. Development of a storage scope display subprogram (SCCPE) and a storage scope plotting program (PLOT). These programs provide for rapid display of graphical information. Subroutine PLOT is compatible with the corresponding hard-copy plotter program so it is useful for rapid checkout of programs using that device
3. Development of storage scope image display programs (DSPF, DSPC). These have greatly reduced the time required to display processed images on the storage scope.
4. Development of an improved text editor for editing computer programs and data files.
5. Development of a driver program for the carousel-projector display.

6. Development of a random number generator for use in generating synthetic subject files for checking operation of software to be developed.

In addition, study of the structure of the Oakland CRIME system was begun. Preliminary investigations indicated that we would be unable to install the whole CRIME system until we increased the size of the memory of the HP2100 computer. Personnel with responsibility for software development were used extensively to support others with less knowledge of the computer operating system.

Forgery Applications

In cases where a photograph of a suspect is available, application of the pattern recognition approach does not depend on the memory ability of a witness. One important area of this type (suggested for use by members of the Houston Police Department) is the forgery area. In a large number of cases they have a forged check and a photograph of the person who cashed the check. The problem is to determine if the individual in the photograph is in their mug file of known forgery offenders including other current unsolved cases.

During the last two months of Phase I, two graduate students from the project spent considerable time with two detectives from the Houston Police Department learning the types of information they have to work with and the procedures they use. They determined several possible approaches which could be used to apply the pattern recognition approach if we can overcome two major problems:

1. Photos of suspects are taken at a variety of angles so you do not have a "straight ahead" view used in mug shots. A procedure must be developed to "rotate" facial images so that images can be compared in mug shot position.
2. Many forgery suspects use "disguises" such as glasses, wigs, etc. Techniques which minimize the effect of disguises must be developed. By the end of Phase I, one of the graduate students was developing a computer program which would test

one approach to the "rotation" problem. This work will be continued in Phase II.

Further Research Phase II

Further Research, Phase II

Phase I of this project was devoted to the development of a "first generation" system. This was a research/development activity using as much as possible the computers and other equipment already owned by the University of Houston. Once this "first generation" system had proved the basic idea would work, the plan for Phase II was to concentrate on planning and designing a "second generation" system which would be appropriate for mug file owners to acquire and install. It was recognized that the agencies who would make this step would probably want a system with capabilities beyond just retrieving a "look alike face" from the mug file. In actual applications there will be additional information about the suspect other than the image of the face; i.e. estimates of height, weight, age, sex, race, type of crime committed, etc. This information should also be used in retrieving suspects from the file. In addition to this requirement, the system should be flexible enough to allow law enforcement agencies to include factors unique to their techniques and filing systems. Our discussions with the Oakland Police Department and Hewlett Packard, who developed the software for the Oakland system, have given us a good understanding of the problems which exist with their system.

Objectives for Phase II

1. To complete the design, build and demonstrate a "second

generation" system which would be appropriate for owners of large mug files. The original proposal included only the design of the system; we now plan to build and demonstrate the system. Because many of the hardware developments have already been achieved, the major requirements to complete this system are software developments.

2. To complete a number of studies which will maximize the effectiveness of the system and minimize the effort and investment required to install this type of system in law enforcement agencies.

Appendices

Appendix HF1

Artist and Technician Credentials

Sketch Artists

Three graduates with B. F. A. degrees from the University of Houston Art Department served as sketch artists. Two of the artists were white males, ages 24 and 28. The third was a white female, age 25. Each artist had previous experience in drawing faces: one had advance training in portrait and the other two had worked with witnesses in doing sketches for the University of Houston Traffic and Security Division.

Although there was no available formal training course for the sketch artists, the work for the Traffic and Security Division did provide insights that led to various interviewing procedures. Other activities that were part of the artist training included having each artist serve as a witness in a sketch situation. This turned out to be a valuable experience since it enabled the artists to be more sensitive to the problems of information exchange in the image generation task. Of course, another dimension of the training was simply doing a number of practice sketches. Each artist drew a minimum of three practice sketches from witnesses' verbal descriptions.

As in the identi-kit training procedures, verbal critiques of each sketch were exchanged among the artists.

Identi-Kit Technicians

Three graduate students in the University of Houston Psychology Program served as identi-kit technicians. Two of the three were white, twenty-six year old, male students in the Cognitive Psychology program. The third was a white, twenty-four year old, female student in the Bio-psychology program.

Formal training in the use of the identi-kit was not identical for all three

technicians. One of the male technicians became a certified identi-kit operator by attending a two and a half day course provided for law enforcement personnel by the Identi-Kit Company. Course content was noted by the attending technician and utilized in developing a training procedure for the remaining two technicians, who did not attend the formal identi-kit course. The resulting training procedure consisted of three phases which preceded formal data collection: (1) a review by the certified technicians of the content of the identi-kit course, (2) the construction of faces from photographs, and (3) the generation of images from verbal descriptions.

In the first phase of training the certified identi-kit technician reviewed the information presented at the identi-kit course. The following points were covered with each of technicians:

(1) The purpose of the face constructed with the kit is to eliminate not to identify. Facial information provided by the kit permits one to eliminate people who could not be candidates.

(2) Procedures utilized in developing the proper setting for the witness were mentioned but not emphasized, since the setting for witnesses in the laboratory were already established.

(3) Construction of identi-kit facial composite begins by asking the witness four basic questions and recording specific responses on a form used with the Identi-Kit. The questions and responses categories include the following:

(a) Approximate height of the suspect? Response categories are: tall, medium, and short. Classification is based on the following table.

	Men	Women
Tall	6'	5'6"
Medium	5'7"-5'11"	5'1"-5'5"
Short	5'6"-	5'-

(b) Build of the suspect? Response categories are heavy, medium, slender, and square.

(c) Age of the suspect? Response categories consist of age groups starting at age 15 and ascending in groups of ten years (15-25, 25-35, 35-45, 45-55, 55- and up)

(d) Hair of the suspect? This question is divided into three parts. The first calls for a description of the hairline across the forehead, the second asks about the color of the hair, and the third about the thickness of the hair. The witness is then asked to look at the card in the identi-kit which contains a large selection of hair styles and select one that is most like the suspect.

(4) The answers to the above four questions guides the technician in producing a basic composite. Each response category for the questions is mapped to a corresponding facial feature or set of facial features in the identi-kit. A card in the identi-kit contains the mappings. The features associated with each description following the questions are selected so that the resulting facial composite is plausible given all responses to the question.

(5) The resulting composite is shown to the witness and the construction of the face proceeds in an interactive fashion. The witness indicates which features of the face are not correct and the manner in which they should be changed. The selection is facilitated by the technician

providing structured alternatives to the witness. From this information, alternative values of the feature are selected which are closer to the witnesses description. Generally the technician should exaggerate in the selection features.

(6) Feature selection is made from a book containing all the features in the identi-kit. The technician avoids showing the features in isolation to the witness. The technician selects the feature based on the witness description. The witness mainly works from the composite. Exceptions include hair selection.

(7) Certain aspects of the face can be influenced during the construction period through the use of the following procedures:

(a) <u>Expression</u> - raise or lower eyebrows,
raise or lower lips

(b) <u>Age</u> - raise or lower chin

(c) For females

eyes - E14 others are E15 and E16

nose - N9, N24

younger nose - N35

older nose - N03

older lips - L30

smiling lips - L08

other female lips - L03, L28, L29

other female eyebrows - D02, D21

(8) When the composite is finished, the witness is asked to rate how closely the composite matches person.

Following the review of course content, each technician served first as as a witness and then as technician in constructing a facial composite.

In the second phase of training all three technicians, including the certified operator, constructed six faces from photographic slides. This procedure was utilized to provide practice in manipulating the foils of the kit and to provide experience with the variable values facial features available in the kit. Verbal critiques of each composite were interchanged among technicians during this phase of training.

In the third training phase each technician constructed three images from witnesses' verbal descriptions. Following the construction from description, each technician constructed a facial composite from viewing the target subject. During this phase, verbal critiques of generated images were interchanged among the technicians.

Appendix HF 2

Various Information Forms Used

in Image Generation Study

EXHIBIT 1

SUBJECT DATA FORM

DATE: ____________

NAME: ______________________ Student Number ________

Target Number: ______________ Subject Number ________

Permanent Address: ______________________________

______________________ Phone Number: ____________

Major: ______________ Classification: FR SO JR SR

Birth date: __________ Height: ________ Weight: ________

Sex: M F

Hair Color: Black Brown Blonde Red Gray/white

Hair Length: Bald Thin Short Medium Long

Eye Color: Brown Blue Green Hazel Other

Complexion: Light, fair Dark/black Freckles or splotchy Pockmarked

Accessories: Glasses______ Moustache ______ Beard ________

Visible scar ________ Sideburns ________ none ________

Peculiarities: Visible scars ________ moles ________ birthmarks ________

Others ________

Build: Light Heavy Medium

Race: white black chicano oriental other

Images: Photograph _____ Witness Description(s) ________ Target Present _____

Image Production Technique: Sketch ______ Identi-Kit ______ Synthesizer ______

Other________

Color Photographs: WG__________ WOG __________

Front Bust ________ __________ __________

Profile Bust ________ __________ __________

Transcript(s):

Comments:

EXHIBIT 2

SUGGESTIVE INTERVIEW PROCEDURE

SKETCH ARTIST INFORMATION

DATE: ______________________

TIME: Start __________ Stop __________

Target No. name ______________________

Witness No. name ______________________

Target Information:

Age: ______________

Build: Slender Medium Heavy

Color of Hair: Blonde, Brown, Black, Red, Gray

Color of Eyes: Blue, Green, Hazel, Brown
Light, Medium, Dark

Complexion: Fair, Tan, Dark
Smooth, Rough, Wrinkled, Facial scars

Accessories: Glasses, moustache, beard, side burns, head gear.

Drawing with target present __________.

Sketch Artist Technician ______________________
Signature

EXHIBIT 3

DATE:

TIME

Start:

Stop:

Subject No. ______

Target No. ______

SUGGESTED INTERROGATION PROCEDURE

IDENTI-KIT - IDMO INFORMATION

RACE

White
Black
Other

SEX

Male
Female

AGE GROUP

UNDER 34
A up to 20
B 21 - 25
C 26 - 30
D 31 34

BETWEEN 35 - 45
E 35 - 40
F 41 - 45

OVER 46
G 46 - 50
H 51 - 55
I 56 - 60
J 61 - 65
K Over 65

HEIGHT

TALL - 6' and Over
MEDIUM - 5'7" - 5' 11"
SHORT - Under 5' 6"

BUILD

Slender
Square
Medium
Heavy

COLOR OF HAIR

Blond or Red
Brown
Black
Grey
Bald
Greying

ODDITY (If any)

Note: ______

______

______

SUPPLEMENTAL INFORMATION

Glasses
Mustache
Beard
Side Burns (large)

Hat or Cap
Mask
Tattoo
Freckles

Wrinkles
Acne
Cripple
Facial Scars

Other:

Confidence Level

IMPORTANT: Record Identi-Kit Code for Future Construction:

Identi-Kit Code: ______

IDMO "324" Jacket No. ______

Identi-Kit Technician ______
Name

Portrait Identi-Kit Code:

EXHIBIT 4

SUBJECT COMMENT SHEET

1. When you viewed the target, what did you do to help you remember his face.

______

2. What parts of the face were easiest to remember?

______

3. What parts of the face were difficult to remember?

______

4. What parts of the face were hard to describe?

______

5. What parts of the face were easiest to describe?

______

6. Have you ever had to describe a persons face before? If yes, why?

______

7. If you have any additional comments or thoughts about your experience in this experiment which you feel to be important, describe them below.

______

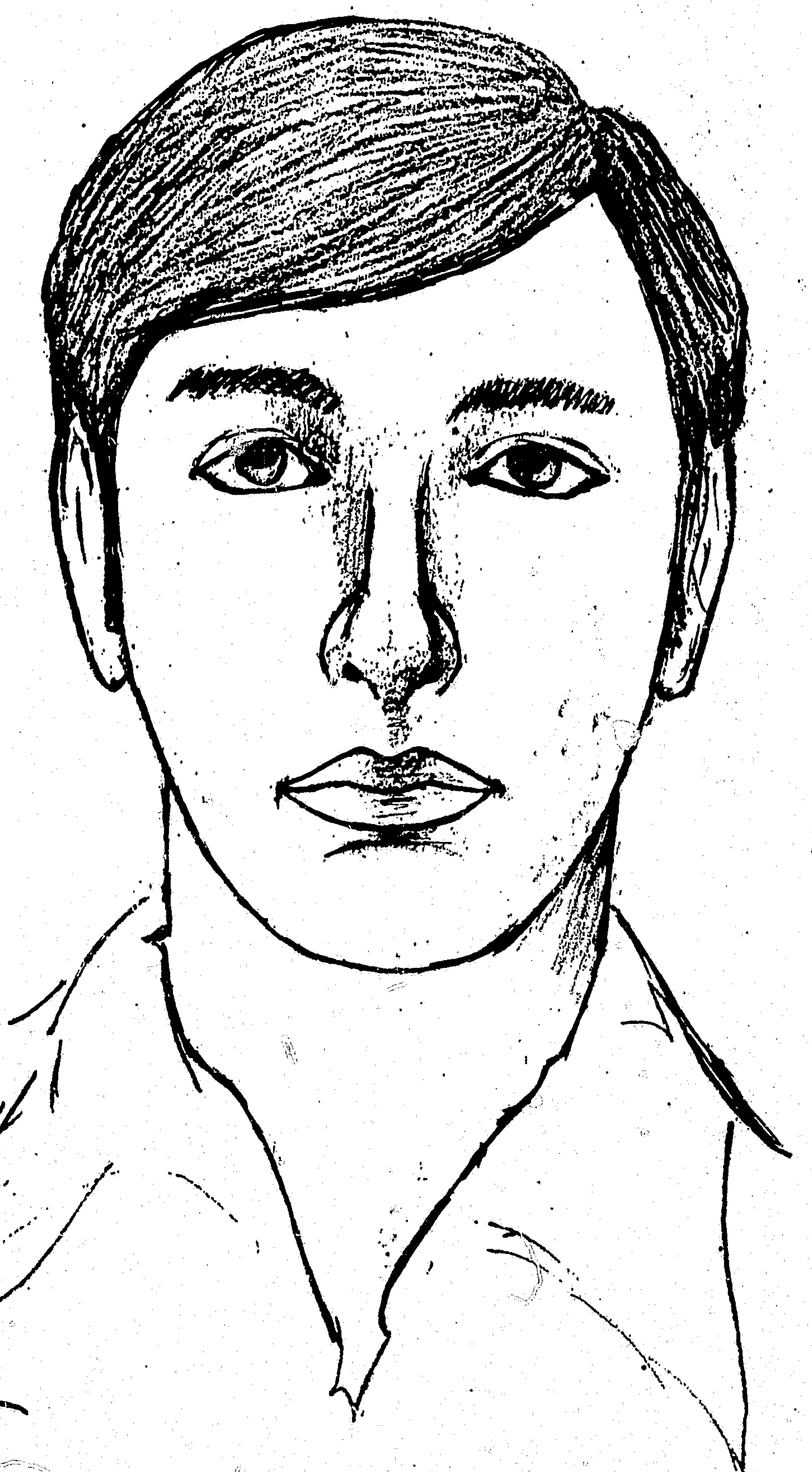
Appendix 3

Sample Photographs and Images from

Image Generation Study

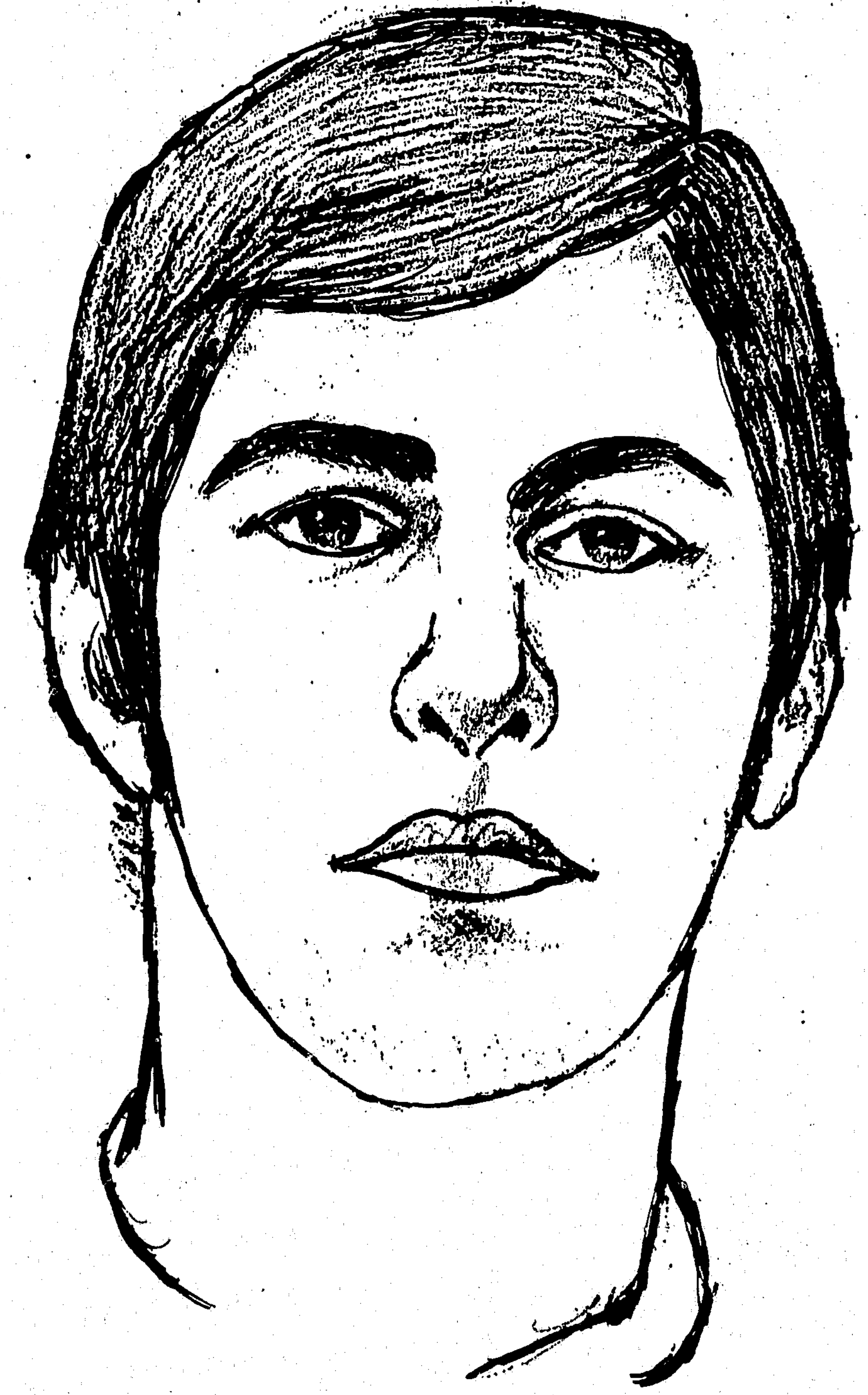
Exhibit 1

Exhibit 2



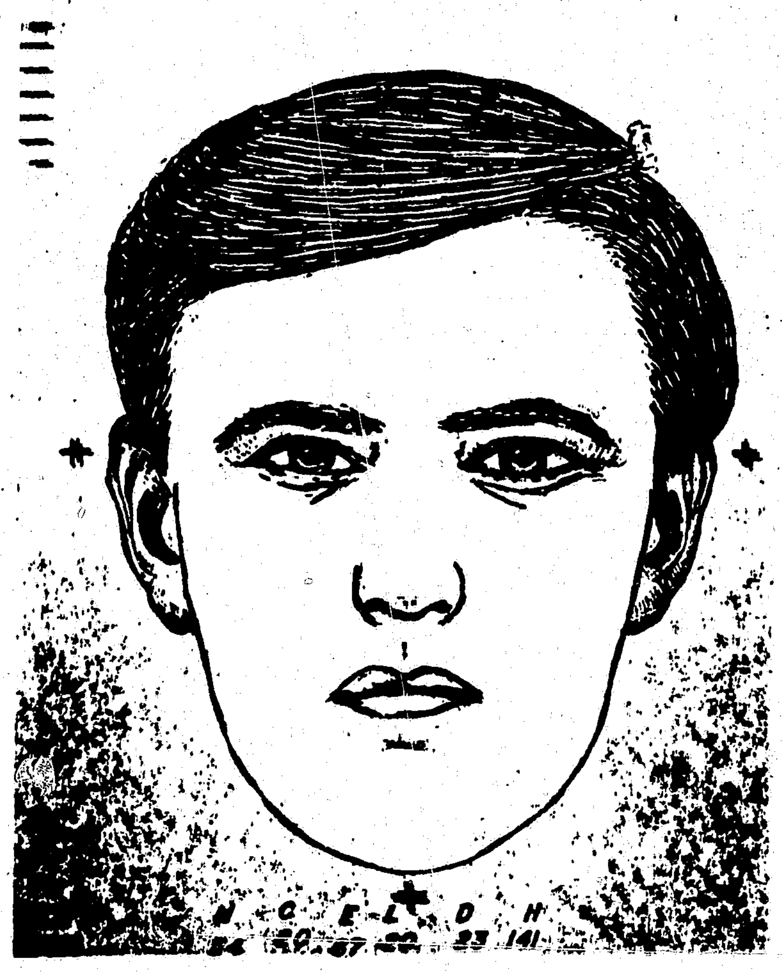
8/14/74
T59
0109
McCoy

Exhibit 3



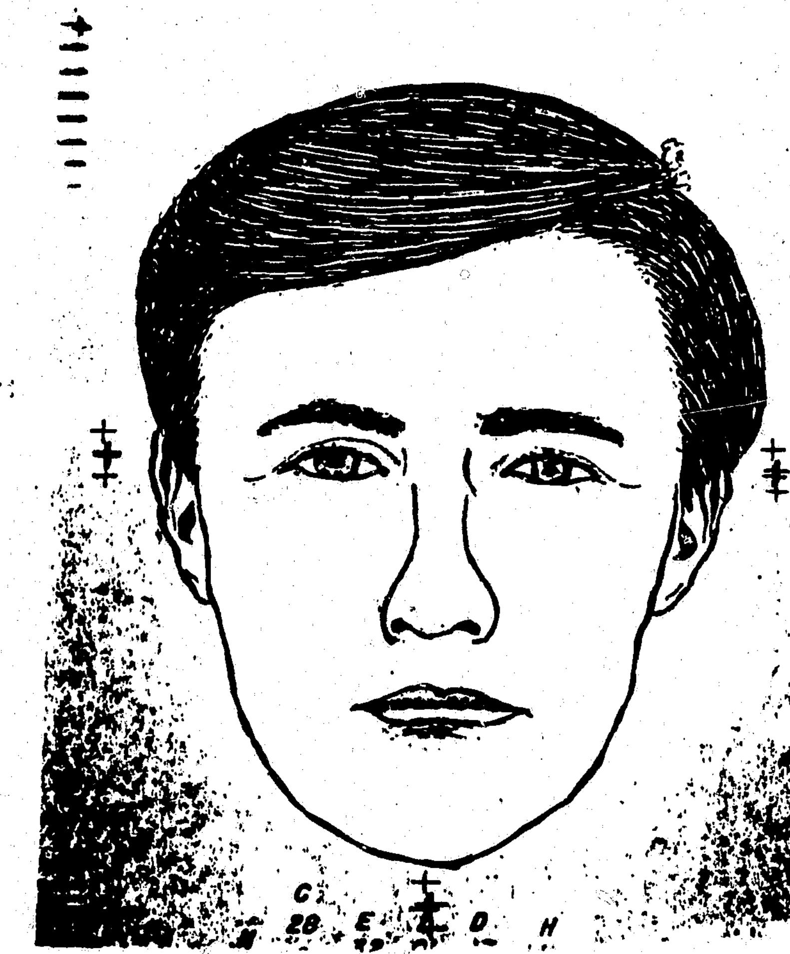
8/14/74
T59
McCoy

Exhibit 4



T59
w110
R.F

Exhibit 5



T59
P
RF

Appendix HF 4

Image Generation Study: Instructions for Subjects and Conversational Mode of Witness-Target Interactions.

Exhibit I

Prototype Instructions to Witness Subjects

(In the following instructions WS1 and WS2 are substituted for the subjects' names)

WS1 and WS2, now that I've finished taking the photographs, we are going to go to the room next door where I will introduce you to another participant in this study. The person you meet is someone you will later attempt to describe for purposes of producing an image of him. The experiment is set up so that you and the person will spend about seven to ten minutes talking with each other. Following this conversation, one of you will work with a sketch artist and the other with an indenti-kit technician. Your task will be to describe from memory the target person you have seen in order to produce a likeness of him.

Exhibit 2

Prototype Instructions to Target Subjects

(In the following instructions TS is substituted for the subject's name)

TS, in a few minutes I will bring two other subjects into this room to meet you. We will spend about seven to ten minutes talking with each other. We use this conversation to give the other subjects an opportunity to see you so they can then describe you from memory. This is the purpose of the study, to see how successfully people can participate in producing an image of someone they have seen. It will help the interaction process go smoothly if you and they can get an easy conversation going.

Exhibit 3

Prototype Introductory Remarks for Witness-Target Conversational Interaction

(In the following statement WS1, WS2 and TS are substituted for the subjects' names)

"WS1 and WS2, I would like you to meet TS. WS1 and WS2, if you will sit opposite TS and me we will take a few minutes for you to get acquainted with TS. As you know (looking at WS1 and WS2), you are going to be working with either a sketch artist or identi-kit technician to develop a facial image of TS. TS, while WS1 and WS2 are giving their descriptions, we will go next door where you can fill out a data form and I will take some pictures of you. We will use one of the photographs as the standard against which we will compare WS1's and WS2's images. In addition to the photographs, TS, we will ask you to pose while our sketch artist and identi-kit technician prepare an image while viewing you."

The above statement was made by E primarily because it created a feeling of mutual participation between the subjects. Following the statement, E would attempt to get a conversation started around the witnesses' and target's activities and interests.

Appendix HF5

Betts and Gordon Imagery Test Forms

THE BETTS QMI VIVIDNESS OF IMAGERY SCALE

Instructions for doing test

The aim of this test is to determine the vividness of your imagery. The items of the test will bring certain images to your mind. You are to rate the vividness of each image by reference to the accompanying rating scale, which is shown at the bottom of the page. For example, if your image is 'vague and dim' you give it a rating of 5. Record your answer in the brackets provided after each item. Just write the appropriate number after each item. Before you turn to the items on the next page, familiarize yourself with the different categories on the rating scale. Throughout the test, refer to the rating scale when judging the vividness of each image. A copy of the rating scale will be printed on each page. Please do not turn to the next page until you have completed the items on the page you are doing, and do not turn back to check on other items you have done. Complete each page before moving on to the next page. Try to do each item separately independent of how you may have done other items.

The image aroused by an item of this test may be:

Perfectly clear and as vivid as the actual experience	Rating 1
Very clear and comparable in vividness to the actual experience	Rating 2
Moderately clear and vivid	Rating 3
Not clear or vivid, but recognizable	Rating 4
Vague and dim	Rating 5
So vague and dim as to be hardly discernible	Rating 6
No image present at all, you only 'knowing' that you are thinking of the object	Rating 7

An example of an item on the test would be one which asked you to consider an image which comes to your mind's eye of a red apple. If your visual image was moderately clear and vivid you would check the rating scale and mark '3' in the brackets as follows:

Item	Rating
5. A red apple	(3)

Now turn to the next page when you have understood these instructions and begin the test.

Think of some relative or friend whom you frequently see, considering carefully the picture that rises before your mind's eye. Classify the images suggested by each of the following questions as indicated by the degrees of clearness and vividness specified on the Rating Scale.

CONTINUED

1 OF 2

Item	Rating
1. The exact contour of face, head, shoulders and body	( )
2. Characteristic poses of head, attitudes of body, etc.	( )
3. The precise carriage, length of step, etc. in walking	( )
4. The different colours worn in some familiar costume	( )

Think of seeing the following, considering carefully the picture which comes before your mind's eye; and classify the image suggested by the following question as indicated by the degree of clearness and vividness specified on the Rating Scale.

5. The sun as it is sinking below the horizon ( )

Rating Scale

The image aroused by an item of this test may be:

Perfectly clear and as vivid as the actual experience	Rating 1
Very clear and comparable in vividness to the actual experience	Rating 2
Moderately clear and vivid	Rating 3
Not clear or vivid, but recognizable	Rating 4
Vague and dim	Rating 5
So vague and dim as to be hardly discernible	Rating 6
No image present at all, you only 'knowing' that you are thinking of the object	Rating 7

Think of each of the following sounds, considering carefully the image which comes to your mind's ear, and classify the images suggested by each of the following questions as indicated by the degrees of clearness and vividness specified on the Rating Scale.

Item	Rating
6. The whistle of a locomotive	( )
7. The honk of an automobile	( )
8. The mewing of a eat	( )
9. The sound of escaping steam	( )
10. The clapping of hands in applause	( )

Rating Scale

The image aroused by an item of this test may be:

Perfectly clear and as vivid as the actual experience	Rating 1
Very clear and comparable in vividness to the actual experience	Rating 2
Moderately clear and vivid	Rating 3

Not clear or vivid, but recognizable	Rating 4
Vague and dim	Rating 5
So vague and dim as to be hardly discernible	Rating 6
No image present at all, you only 'knowing' that you are thinking of the object	Rating 7

Think of 'feeling' or touching each of the following, considering carefully the image which comes to your mind's touch, and classify the images suggested by each of the following questions as indicated by the degrees of clearness and vividness specified on the Rating Scale.

Item	Rating
11. Sand	( )
12. Linen	( )
13. Fur	( )
14. The prick of a pin	( )
15. The warmth of a tepid bath	( )

Rating Scale

The image aroused by an item of this test may be:

Perfectly clear and as vivid as the actual experience	Rating 1
Very clear and comparable in vividness to the actual experience	Rating 2
Moderately clear and vivid	Rating 3
Not clear or vivid, but recognizable	Rating 4
Vague and dim	Rating 5
So vague and dim as to be hardly discernible	Rating 6
No image present at all, you only 'knowing' that you are thinking of the object	Rating 7

Think of performing each of the following acts, considering carefully the image which comes to your mind's arms, legs, lips, etc., and classify the images suggested as indicated by the degree of clearness and vividness specified on the Rating Scale.

Item	Rating
16. Running upstairs	( )
17. Springing across a gutter	( )
18. Drawing a circle on paper	( )

19. Reaching up to a high shelf ( )
20. Kicking something out of your way ( )

Rating Scale

The image aroused by an item of this test may be:

Perfectly clear and as vivid as the actual experience	Rating 1
Very clear and comparable in vividness to the actual experience	Rating 2
Moderately clear and vivid	Rating 3
Not clear or vivid, but recognizable	Rating 4
Vague and dim	Rating 5
So vague and dim as to be hardly discernible	Rating 6
No image present at all, you only 'knowing' that you are thinking of the object	Rating 7

Think of tasting each of the following considering carefully the image which comes to your mind's mouth, and classify the images suggested by each of the following by each of the following questions as indicated by the degrees of clearness and vividness specified on the Rating Scale.

Item	Rating
21. Salt	( )
22. Granulated (white) sugar	( )
23. Oranges	( )
24. Jelly	( )
25. Your favourite soup	( )

Rating Scale

The image aroused by an item of this test may be:

Perfectly clear and as vivid as the actual experience	Rating 1
Very clear and comparable in vividness to the actual experience	Rating 2
Moderately clear and vivid	Rating 3
Not clear or vivid, but recognizable	Rating 4
Vague and dim	Rating 5
So vague and dim as to be hardly discernible	Rating 6
No image present at all, you only 'knowing' that you are thinking of the object	Rating 7

Think of smelling each of the following, considering carefully the image which comes to your mind's nose and classify the images suggested by each of the following questions as indicated by the degrees of clearness and vividness specified on the Rating Scale.

Item	Rating
26. An ill-ventilated room	( )
27. Cooking cabbage	( )
28. Roast beef	( )
29. Fresh paint	( )
30. New leather	( )

Rating Scale

The image aroused by an item of this test may be:

Perfectly clear and as vivid as the actual experience	Rating 1
Very clear and comparable in vividness to the actual experience	Rating 2
Moderately clear and vivid	Rating 3
Not clear or vivid, but recognizable	Rating 4
Vague and dim	Rating 5
So vague and dim as to be hardly discernible	Rating 6
No image present at all, you only 'knowing' that you are thinking of the object	Rating 7

Think of each of the following sensations, considering carefully the image which comes before your mind, and classify the images suggested as indicated by the degrees of clearness and vividness specified on the Rating Scale.

Item	Rating
31. Fatigue	( )
32. Hunger	( )
33. A sore throat	( )
34. Drowsiness	( )
35. Repletion as from a very full meal	( )

The image aroused by an item of this test may be:

Perfectly clear and as vivid as the actual experience	Rating 1
Very clear and comparable in vividness to the actual experience	Rating 2
Moderately clear and vivid	Rating 3

Not clear or vivid, but recognizable	Rating 4
Vague and dim	Rating 5
So vague and dim as to be hardly discernible	Rating 6
No image present at all, you only 'knowing' that you are thinking of the object	Rating 7

THE GORDON TEST OF VISUAL IMAGERY CONTROL

You have just completed a questionnaire that was designed to measure the vividness of different kinds of imagery. In this present questionnaire some additional aspects of your imagery are being studied.

The questions are concerned with the ease with which you can control or manipulate visual images. For some people this task is relatively easy and for others relatively hard. One subject who could not manipulate his imagery easily gave this illustration. He visualized a table, one of whose legs suddenly began to collapse. He then tried to visualize another table with four solid legs, but found it impossible. The image of the first table with its collapsing leg persisted. Another subject reported that when he visualized a table the image was rather vague and dim. He could visualize it briefly but it was difficult to retain by any voluntary effort. In both these illustrations the subjects had difficulty in controlling or manipulating their visual imagery. It is perhaps important to emphasize that these experiences are in no way abnormal and are as often reported as the controllable type of image.

Read each question, then close your eyes while you try to visualize the scene described. Record your answer by underlining 'Yes' 'No' or 'Unsure', whichever is the most appropriate. Remember that your accurate and honest answer to these questions is most important for the validity of this study. If you have any doubts at all regarding the answer to a question, underline 'Unsure'. Please be certain that you answer each of the twelve questions.

1.	Can you see a car standing in the road in front of a house?	Yes	No	Unsure
2.	Can you see it in colour?	Yes	No	Unsure
3.	Can you now see it in a different colour?	Yes	No	Unsure
4.	Can you now see the same car lying upside down?	Yes	No	Unsure
5.	Can you now see the same car back on its four wheels again?	Yes	No	Unsure
6.	Can you see the car running along the road?	Yes	No	Unsure
7.	Can you see it climb up a very steep hill?	Yes	No	Unsure
8.	Can you see it climb over the top?	Yes	No	Unsure
9.	Can you see it get out of control and crash through a house?	Yes	No	Unsure
10.	Can you now see the same car running along the road with a handsome couple inside?	Yes	No	Unsure
11.	Can you see the car cross a bridge and fall over the side into the stream below?	Yes	No	Unsure
12.	Can you see the car all old and dismantled in a car-cemetery?	Yes	No	Unsure

Appendix HF6

Accessories Study

Facial Recognition: Effects of Changing Accessories[1]

KENNETH R. LAUGHERY and RICHARD H. FOWLER, University of Houston

A facial recognition study examined effects of accessories changes between initial exposure to a target person and that person's appearance in a recognition task. Three accessories were manipulated; glasses and beards (present or absent) and hair style (long or short). Changes in both directions had marked negative effects upon recognition, with hit rates dropping as much as 42 percent. The various accessories had differential effects; glasses producing the smallest decrement and beards the largest. False positive errors were also increased by accessory changes. The results have implications for criminal identification systems.

INTRODUCTION

Two earlier papers (Laughery, Alexander and Lane, 1971; Laughery, Fessler, Lenorovitz and Yoblick, 1974) reported a series of experiments exploring the effects of several task variables on facial recognition. The paradigm in these studies simulated a situation in which a witness who has seen a criminal attempts to identify that person's picture in a set of alternatives. While a number of variables were explored, two in particular strongly affected recognition: the more decoys (or distractors) preceding the target, the poorer the performance; and the more similar the decoys were to the target, the poorer the performance.

The results of these experiments were related to the design of criminal identification systems. An important set of task variables in such a system that have not been explored to date, however, concerns differences between the target's appearance in the initial exposure and his appearance in the recognition task. In all of the earlier studies the target's appearance was basically the same in the two instances. In the real-world of criminal identification there are frequently changes in appearance. The study reported here explored one class of changes; namely, differences in accessories. Accessories refer to parts of the facial stimulus that are not permanent and are relatively easy to modify. Examples would be beards, moustaches, glasses, hair styles and cosmetics. This experiment dealt specifically with three of these; glasses, beards and hair styles.

METHOD

Subjects. The Ss were 480 undergraduate students enrolled in

introductory psychology at the University of Houston. Class credit was given for participation in the study.

Task. The task in this experiment was essentially the same as that reported in Laughery, Alexander and Lane (1971) and Laughery, Fessler, Lenorovitz and Yoblick (1974). Ss first viewed four sequentially presented slides of the target person in different candid positions. The Ss task was to indicate, using a 6-point scale (definitely yes, probably yes, possibly yes, possibly no, probably no, definitely no), whether each picture in a subsequent, sequentially presented test series of slides was or was not the target. The slides were projected so as to be approximately life size on the screen. The target's picture appeared only once in the test series.

Design. The design of the experiment was a 3 x 4 x 4 factorial with all factors manipulated as between-S variables. The conditions of the first variable, accessory, were beard, hair style and glasses. The second variable was the view-search accessory relationship. More specifically, this variable refers to an accessory change between the target's appearance in the initial exposure and his appearance in the search series. The levels of this variable were defined by the accessory being same or different and the actual condition of the accessory. In the case of the beard and glasses accessories, the change related to the presence or absence of the accessory. For hair style the change was long versus short hair. Perhaps the four levels of this view-search variable are better understood by noting the specific view-search relationships. If we think of "with" as referring to the presence of the accessory (or long hair), and "without" as the absence of the accessory (or short hair), then the four conditions of same or different for each of the three accessories were with-with, with-without, without-with, and without-without.

The third variable, target, consisted of four different people, all white males, whose pictures were used as targets.

A total of 10 Ss were run in each of the 48 experimental conditions.

Materials. The people recruited to be target persons were all clean shaven and had a long hair style. A make-up artist prepared the targets for the different accessory conditions. A short wig was used to effect the hairstyle change. The beards were full and included a moustache. The glasses, of course, were simply put on or off. In this manner a full set of photographs, including candid and posed, were taken for each target with each accessory condition. Ten separate targets were made up and their photographs taken. From these 10 four were selected for the study. The selection criteria were concerned primarily with how natural the makeup appeared. Figure 1 shows one of the targets used in the study with the different accessory conditions.

(See Figure 1)

In this experiment the accessories were manipulated independently; that is, no interactions were considered. Putting it another way, in manipulating the presence or absence of an accessory, only 1 accessory was changed. For example, when the target appeared with a beard, he did not wear glasses and appeared with a short hair style. Similarly, when the target appeared with a long hair style, he was clean-shaven and did not wear glasses.

The test series consisted of 74 decoys and the target, all appearing in front, bust views. The decoys were all white males ranging in age from 18 to 28. Half of the decoys in the test series consisted of decoys without glasses, without beards and with short hair. The appearance of the remaining decoys depended upon the accessory condition. If the condition concerned beards, then the remaining 37 slides contained pictures of men with beards. Similarly, if the condition concerned hair or glasses, the remaining pictures contained long hair or glasses respectively.

The order of the decoys was random with the constraint that no more than 4 consecutive decoys were of the same type with respect to presence or absence of the accessory. The physical parameters of all slides were constant (sharpness, scale, lighting, etc.).

The candid position slides showed the target person in positions ranging from left to right side, full length, and bust views. The candid positions were selected from a larger set of photographs of the target with an effort to select those which seemed least posed.

Apparatus. The apparatus consisted of a Kodak Carousel AV 900 projector with a 4 to 6 in., F3.5 Zoom Ektamar Lens and a Da-Lite projection screen.

Procedure. The Ss in each of the 48 experimental conditions were run as a group. Five Ss were seated at each of two long tables, one behind the other, in a normal size classroom. The screen was located at the front center of the room at a height slightly above the seated Ss. The tables were 7.0 and 12.0 feet from the screen. The projector was located at the rear center of the room. The room was darkened to insure good vision of the slides, but with sufficient light to read and mark the answer sheet.

The instructions were presented in two parts. The first part made clear that the Ss would later be looking for a picture of a person whom they were about to see. Following the presentation of the 4 candid photographs of the target for 10 seconds each, the Ss were given the second part of the instructions. This part included details about the use of the answer sheet and a statement that the target might appear in the test series several times, only once, or not at all. In fact, the target appeared just once, in position 69. Presentation of the second part of the instructions required 4 minutes and the test series followed immediately.

During the search sequence, each slide was projected on the screen for seven seconds with two seconds between slides - during which the Ss recorded their responses on answer sheets.

Any S who knew the target person was given credit for participation and excused from the experiment. The Ss were asked to indicate on their answer sheets if they knew any of the decoys. There was a negligible number of responses indicating any S knew a decoy face.

RESULTS

The Yes and No responses to the target picture in the test series are referred to as hits and misses. Similarly, the Yes and No responses to the decoys are referred to as false alarms and correct rejections. For a given S the hit-miss (H-M) score could be a single value from 1 to 6. A score of 6 indicates that the S responded definitely yes when the target appeared, 5 was probably yes, and so on, with a score of 1 indicating a response of definitely no. Two false alarm-correct rejection

(FA-CR) scores were computed for each S. One considered responses to decoys with all accessories absent; the other considered responses to decoys with the accessory present.

Two analyses were carried out on the results. The first was an analysis of varience on the H-M scores. The mean H-M scores for the 12 treatment conditions (collapsed across targets) are displayed in Table 1. The view-search factor had a significant effect, $F(3,432) = 30.31$, $p < .01$, with performance better in the unchanged conditions than in the changed conditions. A significant view-search by accessory interaction $F(6,432) = 2.76$, $p < .025$, reflects differential view-search effects depending on which accessory was changed. The order of greatest to least performance decrement was beard, hair style and glasses. Although the H-M scores were used in the variance analysis, it is helpful in understanding the data to note the percentage of Ss who had a hit (marked a 4, 5 or 6 when the target appeared). These percentages are shown in Table 2, and obviously reflect the effects revealed in the analysis of variance.

(See Table 1)

(See Table 2)

Two interactions involving the targets were also significant: accessory, $F(6,432) = 2.63$, $p < .025$, and view-search condition, $F(9,432) = 2.63$, $p < .01$. Although the interpretation of these results probably lies with idiosyncracies of the target persons, the exact nature of that interpretation is neither evident nor particularly interesting.

The second analysis was based on the FA-CR scores. In computing the two FA-CR scores only those decoys appearing before the target were considered. The first FA-CR score was the S's mean response to the decoys with all accessories absent. The second FA-CR score was the mean response with the accessory present. The accessory present corresponded to the accessory manipulated in the view-search condition of the target. The mean FA-CR scores for each condition are shown in Table 3. The analysis of variance carried out on these data considered viewing condition only in terms of the two initial viewing conditions of the target (accessory present or absent). Decoy was significant, $F(1,456) = 36.6$, $p<.01$, with a higher FA-CR score for decoys with the accessory. The effect of accessory was significant, $F(2,456) = 9.34$, $p<.01$. Performance was poorest (higher scores) for hair style, best for beard, and glasses was intermediate. The decoy by viewing condition interaction reached significance, $F(1,456) = 12.94$, $p<.01$, and indicated the difference between the decoy with and without the accessory was less when the target initially appeared without the accessory than when he initially appeared with it. The significant interaction between decoy and accessory, $F(2,456) = 8.90$, $p<.01$, is the result of a small difference in FA-CR scores for decoys with accessory and without accessory in the beard condition. This is contrasted with larger differences in the case of glasses and still larger differences for hair styles. Finally, the viewing condition by accessory interaction was significant, $F(2,456) = 4.43$, $p<.025$. With glasses and beard, initially viewing the target with the accessory resulted in higher FA-CR scores than when the target was initially viewed without the accessory. However, for hair

style the reverse was true - higher FA-CR scores occurred when the target initially appeared without the accessory.

(See Table 3)

DISCUSSION

In general, the results of this study are consistent with expectations. When a facial accessory change occurs between the initial encounter and the later recognition task, the probability of a correct identification is greatly reduced. In some cases, the probability of a hit is lowered as much as 42%. A point to be noted about these results is that performance is decremented by a change in either direction; that is, when the accessory is added or when it is deleted. Furthermore, the magnitude of the decrement is roughly equal with the two types of changes.

The significant interaction between the view-search and accessory variables makes sense in terms of the amount of change produced in the facial stimulus by adding or subtracting the various accessories. Glasses change a relatively small part of the face. Also, glasses are transparent and some information about the eyes is available and potentially useful when they are present. While a change in hair style does not typically affect the availability of information about other facial features, hair alterations probably produce significant effects because hair itself is an important feature or source of information in the recognition task (Lenorovitz, 1972).

Beards (including moustaches) result in major changes in facial appearance. Information about several features (e.g., chin, jaw line and mouth) is altered or concealed when a beard is added. When the beard is present during the initial exposure of the target, information relevant to later identification is simply not available. Indeed, the beard itself may be processed as relevant information; a possibility supported by the fact that the with-with beard condition results in the best identification performance in the study.

The FA-CR scores reflect the errors made by subjects on the decoy pictures; the false positives corresponding to situations where a wrong person is identified as the target. The results, in general, make sense. The failure of the decoy and viewing condition variables to have an effect when the accessory was a beard, is probably due to the distinctiveness of the various beards. This notion is supported by the low FA-CR scores in the beard conditions. Errors when the accessory was hair showed more mistakes on decoys with long hair, regardless of the targets initial hair condition. It may be that long hair is simply more confusing. When the accessory was glasses and the target initially appeared without them, the errors on decoys with or without glasses were no different. A possible explanation is that Ss were not using information about the eyes, or if they were, it was still available with glasses present. The significant decoy effect when the target initially wore glasses, may be the result of Ss looking for a target wearing them.

Overall, the results of this study have important implications for criminal identification systems. When a criminal's appearance has been changed as a result of accessory differences between initial exposure and

the mug file, lineup or other search procedure, the probability of a correct identification is lowered and false positives may be increased. Judicial procedures must take these facts into account in evaluating evidence based upon recognition by a witness.

It seems reasonable to assume that procedures could be developed which would permit an identification system to deal more effectively with accessory changes. For example, it should be possible to add or change accessories on pictures in a mug file. Such changes are well within the current technology of computerized systems. Of course, the legality of such procedures may be questioned; however, such issues are beyond the scope of this paper.

REFERENCES

Laughery, K. R., Alexander, J. F. and 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, 55, 477-483.

Laughery, K. R., Fessler, P. K., Lenorovitz, D. R. and Yoblick, D. A. Time delay and similarity effects in facial recognition. Journal of Applied Psychology, 1974, 59, 490-496.

Lenorovitz, D. R. The discrimination of similarities and differences in facial appearance: A multidimensional scaling approach. Unpublished Masters Thesis, State University of New York at Buffalo, 1972.

Footnotes

1. Prepared under grant No. 74-NI-99-0023-G from the National Institute of Law Enforcement and Criminal Justice, Law Enforcement Assistance Administration, Department of Justice.

TABLE 1

Mean Hit-Miss Scores

	Unchanged		Changed	
Accessory	With-with	Without-without	With-without	Without-with
Glasses	4.85	5.45	4.08	4.63
Hair Style	5.35	5.30	3.30	4.13
Beard	5.50	5.10	3.28	3.23

TABLE 2

Percent Hits

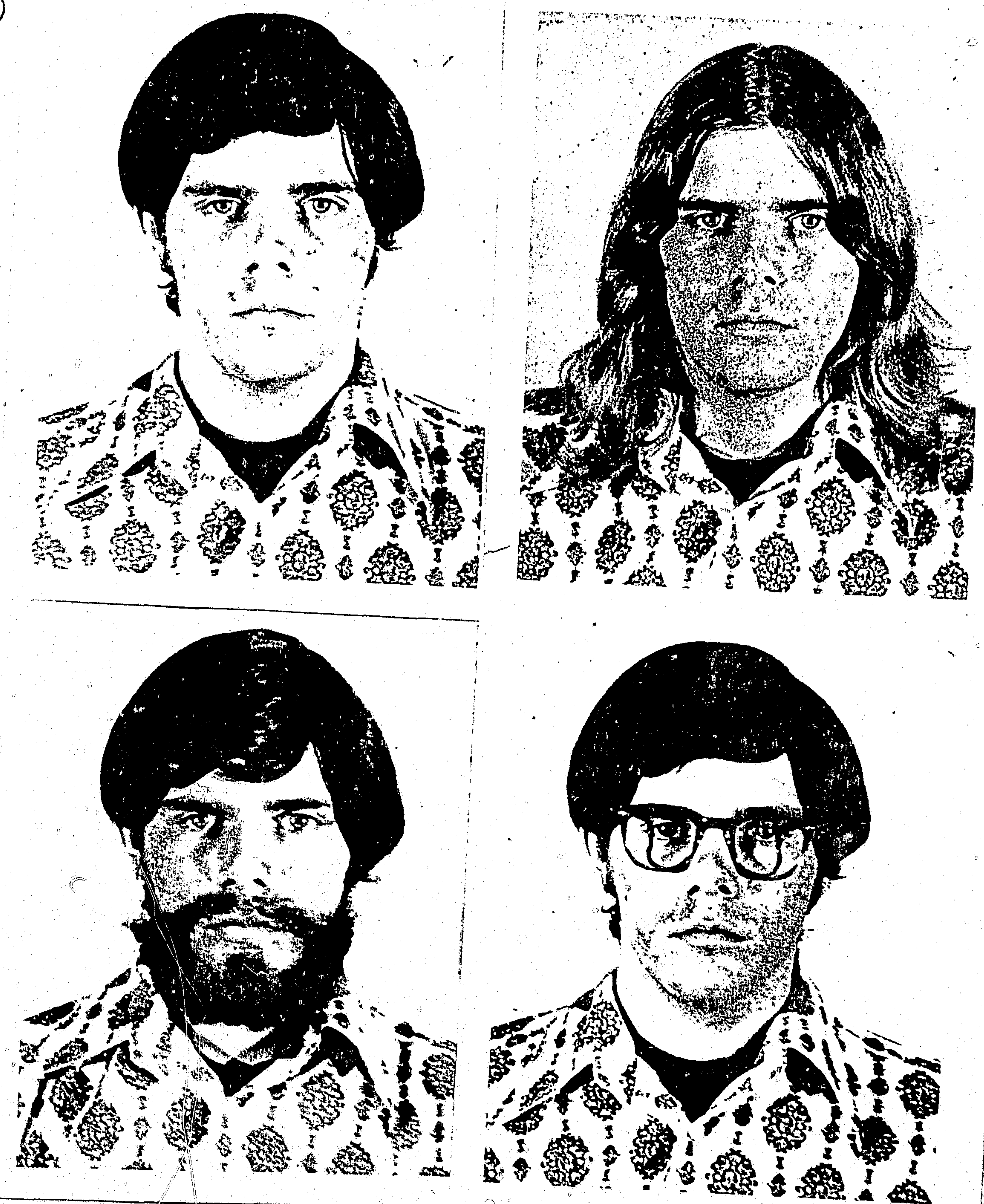
	Unchanged		Changed	
Accessory	With-with	Without-without	With-without	Without-with
Glasses	82.5	92.5	65.0	77.5
Hair Style	90.0	87.5	47.5	67.5
Beard	92.5	82.5	50.0	52.5

TABLE 3

Mean False Alarm-Correct Rejection (FA-CR) Scores

Accessory	Target with: Decoy with	Target with: Decoy without	Target without: Decoy with	Target without: Decoy without
Glasses	1.45	1.30	1.25	1.24
Beard	1.27	1.23	1.16	1.19
Hair	1.43	1.28	1.59	1.48

Figure 1. Example of Target with and without different accessories.



Facial Recognition: Effects of Changing Accessories[1]

Kenneth R. Laughery and Richard H. Fowler

University of Houston

Abstract

A facial recognition study examined effects of accessories changes between initial exposure to a target person and that person's appearance in a recognition task. Three accessories were manipulated; glasses and beards (present or absent) and hair style (long or short). Changes in both directions had marked negative effects upon recognition, with hit rates droping as much as 42 percent. The various accessories had differential effects; glasses producing the smallest decrement and beards the largest. False positive errors were also increased by accessory changes. The results have implications for criminal identification systems.

Facial Recognition: Effects of Changing Accessories[1]

Kenneth R. Laughery and Richard H. Fowler

University of Houston

Two earlier papers (Laughery, Alexander and Lane, 1971; Laughery, Fessler, Lenorovitz and Yoblick, 1974) reported a series of experiments exploring the effects of several task variables on facial recognition. The paradigm in these studies simulated a situation in which a witness who has seen a criminal attempts to identify that person's picture in a set of alternatives. While a number of variables were explored, two in particular strongly affected recognition: the more decoys (or distractors) preceding the target, the poorer the performance; and the more similar the decoys were to the target, the poorer the performance.

The results of these experiments were related to the design of criminal identification systems. An important set of task variables in such a system that have not been explored to date, however, concerns differences between the target's appearance in the initial exposure and his appearance in the recognition task. In all of the earlier studies the target's appearance was basically the same in the two instances. In the real-world of criminal identification there are frequently changes in appearance. The study reported here explored one class of changes; namely, differences in accessories. Accessories refer to parts of the facial stimulus that are not permanent and are relatively easy to modify. Examples would be beards, moustaches, glasses, hair styles and cosmetics. This experiment dealt specifically with three of these; glasses, beards and hair styles.

Method

Subjects. The Ss were 480 undergraduate students enrolled in introductory psychology at the University of Houston. Class credit was given for participation in the study.

Task. The task in this experiment was essentially the same as those reported in Laughery, Alexander and Lane (1971) and Laughery, Fessler, Lenorovitz and Yoblick (1974). Ss first viewed four sequentially presented slides of the target person in different candid positions. The Ss task was to indicate, using a 6-point scale (definitely yes, probably yes, possibly yes, possibly no, probably no, definitely no), whether each picture in a subsequent, sequentially presented test series of slides was or was not the target. The slides were projected so as to be approximately life size on the screen. The target's picture appeared only once in the test series.

Design. The design of the experiment was a 3 x 4 x 4 factorial with all factors manipulated as between-S variables. The conditions of the first variable, accessory, were beard, hair style, glasses. The second variable was the view-search accessory relationship. More specifically, this variable refers to an accessory change between the targets appearance in the initial exposure and his appearance in the search series. The levels of this variable were defined by the accessory being same or different and the actual condition of the accessory. In the case of the beard and glasses accessories, the change related to the presence or absence of the accessory. For hair style the change was long versus short hair. Perhaps the four levels of this view-search variable, are better understood by noting the specific view-search relationships. If we think of "with" as referring to the presence of the accessory (or long hair), and "without" as the absence of the accessory (or short hair), then the four conditions of same or different for each of the three accessories were with-with, with-without, without-with, and without-without.

The third variable, target, consisted of four different people, all white males, whose pictures were used as targets.

A total of 10 Ss were run in each of the 48 experimental conditions.

Materials. The people recruited to be target persons were all clean shaven and had a long hair style. A make-up artist prepared the targets for the different accessory conditions. Ten separate targets were made up and their photographs taken. From these 10 four were selected for the study. The selection criteria was concerned primarily with how natural the makeup appeared.

A short wig was used to effect the hairstyle change. The beards were full and included a moustache. The glasses, of course, were simply put on or off. In this manner, a full set of photographs, including casual and posed, were taken for each target with each accessory condition.

In this experiment the accessories were manipulated independently; that is, no interactions were considered. Putting it another way, in manipulating the presence or absence of an accessory, only 1 accessory was changed. For example, when the target appeared with a beard, he did not wear glasses and appeared with a short hair style. Similarly, when the target appeared with a long hair style, he was clean-shaven and did not wear glasses.

The test series consisted of 74 decoys and the target, all appearing in front, bust views. The decoys were all white males ranging in age from 18 to 28. Half of the decoys in the test series consisted of decoys without glasses, without beards and with short hair. The appearance of the remaining decoys depended upon the accessory condition. If the condition concerned beards, then the remaining 37 slides contained pictures of men with beards. Similarly, if the condition concerned hair or glasses, the remaining pictures contained long hair or glasses respectively.

The order of the decoys was random with the constraint that no more than 4 consecutive decoys were of the same type with respect to presence or absence of the accessory. The physical parameters of all slides were constant (sharpness, scale, lighting, etc.).

The candid position slides showed the target person in positions ranging from left to right side, full length, and bust views. The candid positions were selected from a larger set of photographs of the target with an effort to select those which seemed least posed.

Apparatus. The apparatus consisted of a Kodak Carousel AV 900 projector with a 4 to 6 in., F3.5 Zoom Ektamar Lens and a Da-Lite projection screen.

Procedure. The Ss in each of the 48 experimental conditions were run as a group. Five Ss were seated at each of two long tables, one behind the other, in a normal size classroom. The screen was located at the front center of the room at a height slightly above the seated Ss. The tables were 7.0 and 12.0 feet from the screen. The projector was located at the rear center of the room. The room was darkened to insure good vision of the slides, but with sufficient light to read and mark the answer sheet.

The instructions were presented in two parts. The first part made clear that the Ss would later be looking for a picture of a person whom they were about to see. Following the presentation of the 4 candid photographs of the target for 10 seconds each, the Ss were given the second part of the instructions. This part included details about the use of the answer sheet and a statement that the target might appear in the test series several times, only once, or not at all. In fact, the target appeared just once, in position 68. Presentation of the second part of the instructions required 4 minutes and the test series followed immediately.

During the search sequence, each slide was projected on the screen for seven seconds with two seconds between slides - during which the Ss recorded their responses on answer sheets.

Any S who knew the target person was given credit for participation and excused from the experiment. The Ss were asked to indicate on their answer sheets if they knew any of the decoys. There was a negligible number of responses indicating any S knew a decoy face.

Results

The Yes and No responses to the target picture in the test series are referred to as hits and misses. Similarly, the Yes and No responses to the decoys are referred to as false alarms and correct rejections. For a given S the hit-miss (H-M) score could be a single value from 1 to 6. A score of 6 indicates that the S responded definitely yes when the target appeared, 5 was probably yes, and so on, with a score of 1 indicating a response of definitely no. Two false alarm-correct rejection (FA-CR) scores were computed for each S. One considered responses to decoys with all accessories absent: the other considered responses to decoys with the accessory present.

Two analyses were carried out on the results. The first was an analysis of variance on the H-M scores. The mean H-M scores for the 12 treatment conditions (collapsed across targets) are displayed in Table 1. The view-search factor had a significant effect, $F(3,432)=30.31$, $p<.01$, with performance better in the unchanged conditions than in the changed conditions. A significant view-search by acessory interaction $F(6,432)= 2.76$, $p<.025$, reflects differential view-search effects depending on which accessory was changed. The order of greatest to least performance decrement was beard, hair style and glasses. Although the H-M scores were used in the variance analysis, it is helpful in understanding the data to note the percentage of Ss who had a hit (marked a 4, 5 or 6 when the target appeared). These percentages are shown in Table 2, and obviously reflect the effects revealed in the analysis of variance.

Two interactions involving the targets were also significant: accessory, $F(6,432) = 2.63$, $p<.025$, and view-search condition, $F(9,432) = 2.63$, $p<.01$. Although the interpretation of these results probably lies with idiosyncracies of the target persons, the exact nature of that interpretation is neither evident nor particularly interesting.

The second analysis was based on the FA-CR scores. In computing the two FA-CR scores only those decoys appearing before the target were considered. The first FA-CR score was the S's mean response to the decoys with all accessories absent. The second FA-CR score was the mean response with the accessory present. The accessory present corresponded to the accessory manipulated in the view-search condition of the target. The mean FA-CR scores for each condition are shown in Table 3. The analysis of variance carried out on these data considered viewing condition only in terms of the two initial viewing conditions of the target (accessory present or absent). Decoy was significant, $F(1,456) = 36.6$, $p<.01$, with a higher FA-CR score for decoys with the accessory. The effect of accessory was significant, $F(2,456) = 9.34$, $p<.01$. Performance was poorest (higher scores) for hair style, best for beard, and glasses was intermediate. The decoy by viewing condition interaction reached significance, $F(1,456) = 12.94$, $p<.01$, and indicated the difference between the decoy with and without the accessory was less when the target initially appeared without the accessory than when he initially appeared with it. The significant interaction between decoy and accessory, $F(2,456) = 8.90$, $p<.01$, is the result of a small difference

in FA-CR scores for decoys with accessory and without accessory in the beard condition. This is contrasted with larger differences in the case of glasses and still larger differences for hair styles. Finally, the viewing condition by accessory interaction was significant, $F(2,456) = 4.43$, $p<.025$. With glasses and beard, initially viewing the target with the accessory resulted in higher FA-CR scores than when the target was initially viewed without the accessory. However, for hair style the reverse was true - higher FA-CR scores occurred when the target initially appeared without the accessory.

Discussion

In general, the results of this study are consistent with expectations. When a facial accessory change occurs between the initial encounter and the later recognition task, the probability of a correct identification is greatly reduced. In some cases, the probability of a hit is lowered as much as 40%. A point to be noted about these results is that performance is decremented by a change in either direction; that is, when the accessory is added or when it is deleted. Furthermore, the magnitude of the decrement is roughly equal with the two types of changes.

The significant interaction between the view-search and accessory variables makes sense in terms of the amount of change produced in the facial stimulus by adding or subtracting the various accessories. Glasses change a relatively small part of the face. Also, glasses are transparent and some information about the eyes is available and potentially useful when they are present. While a change in hair style does not typically affect the availability of information about other facial features, hair alterations probably produce significant effects because hair itself is an important feature or source of information in the recognition task (Lenorovitz, 1972).

Beards (including moustaches) result in major changes in facial appearance. Information about several features (e.g., chin, jaw line and mouth) is altered or concealed when a beard is added. When the beard is present during the initial exposure of the target, information relevant to later identification is simply not available. Indeed, the beard itself may be processed as relevant information; a possibility supported by the fact that the with-with beard condition results in the best identification performance in the study.

The FA-CR scores reflect the errors made by subjects on the decoy pictures; the false positives corresponding to situations where a wrong person is identified as the target. The results, in general, make sense. The failure of the decoy and viewing condition variables to have an effect when the accessory was a beard, is probably due to the distinctiveness of the various beards. This notion is supported by the low FA-CR scores in the beard conditions. Errors when the accessory was hair showed more mistakes on decoys with long hair, regardless of the targets initial hair condition. It may be that long hair is simply more confusing. When the accessory was glasses and the target initially appeared without them, the errors on decoys with or without glasses were no different. Most likely, Ss were not using information about the eyes, or if they were, it was still available with glasses present. The significant decoy effect when the target initially wore glasses, may be the result of Ss looking for a target wearing them.

Overall, the results of this study have important implications for criminal identification systems. When a criminal's appearance has been changed as a result of accessory differences between initial exposure and the mug file, lineup or other search procedure, the probability of a correct identification is lowered

and false positives are increased. Judicial procedures must take these facts into account in evaluating evidence based upon recognition by a witness.

It seems reasonable to assume that procedures could be developed which would permit an identification system to deal more effectively with accessory changes. For example, it should be possible to add or change accessories on pictures in a mug file. Such changes are well within the current technology of computerized systems. Of course, the legality of such procedures may be questioned; however, such issues are beyond the scope of this paper.

References

1. Laughery, K. R., Alexander, J. F. and 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, 55, 477-483.

2. Laughery, K. R., Fessler, P. K., Lenorovitz, D. R. and Yoblick, D.A. Time delay and similarity effects in facial recognition. Journal of Applied Psychology, 1974, 59, 490-496.

3. Lenorovitz, D. A. The discrimination of similarities and differences in facial appearance: A multidimensional scaling approach. State University of New York at Buffalo, Unpublished Masters Thesis, 1972.

Table 1

Mean Hit-Miss Scores

	Unchanged		Changed	
Accessory	With-with	Without-without	With-without	Without-with
Glasses	4.85	5.45	4.08	4.63
Hair Style	5.35	5.30	3.30	4.13
Beard	5.50	5.10	3.28	3.23

Table 2

Percent Hits

Accessory	Unchanged: With-with	Unchanged: Without-without	Changed: With-without	Changed: Without-with
Glasses	82.5	92.5	65.0	77.5
Hair style	90.0	87.5	47.5	67.5
Beard	92.5	82.5	50.0	52.5

Table 3

Mean False Alarm-Correct Rejection (FA-CR) Scores

	Target with		Target without	
	Decoy with	Decoy without	Decoy with	Decoy without
Glasses	1.45	1.30	1.25	1.24
Beard	1.27	1.23	1.16	1.19
Hair	1.43	1.28	1.59	1.48

Footnotes

1. Prepared under grant No. 74-NI-99-0023-G from the National Institute of Law Enforcement and Criminal Justice, Law Enforcement Assistance Administration, Department of Justice.

END

END