University of Houston

MUG FILE PROJECT REPORT NUMBER UHMUG- 3

Factors Affecting Facial Recognition

K. R. Laughery and R. H. Fowler.

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

MUG FILE PROJECT REPORTS

UHMUG-1 Summary report for a Research Project *"A Man-Computer System for Solution of the Mug File Problem"*.
B. T. Rhodes, K. R. Laughery, G. M. Batten, and J. D. Bargainer.

UHMUG-2 *An Analysis of Procedures for Generating Facial Images*
K. R. Laughery, G. C. Duval, and R. H. Fowler.

UHMUG-3 *Factors Affecting Facial Recognition*
K. R. Laughery and R. H. Fowler

UHMUG-4 *The Minolta Montage Synthesizer as a Facial Image Generating Device*
F. H. Duncan and K. R. Laughery

UHMUG-5 *An Analysis of Strategies in Remembering and Generating Faces*
G. C. Duval

UHMUG-6 Data Base No. 1 - *Sketches and Identi-Kit Composites*

UHMUG-7 Data Base No. 2 - *Transcripts of Artist/Technician and Witness Interaction*

UHMUG-8 Data Base No. 3 - *Adjective Descriptors Used in Generating Sketches and Identi-Kit Composites*

UHMUG-9 Data Base No. 4 - *Miscellaneous Data from Sketch and Identi-Kit Generation*

UHMUG-10 *Support Hardware for Image Analysis Techniques Applied to the Mug File Program*
J. D. Bargainer

UHMUG-11 *Forgery Application of a Pattern Recognition Algorithm for Facial Images*
B. T. Rhodes and K. Prasertchuang

UHMUG-12 *An Evaluation of the UHMFS Facial Image Pattern Recognition Algorithms*
B. T. Rhodes and C. R. Walters

UHMUG-13 *FORTRAN Subroutines for the Pattern Recognition Algorithm Designed to Find "Look-Alikes" in a Mug File*
K. Sumney

UHMUG-14 *A Computer Simulation of the Minolta Montage Synthesizer*
G. W. Batten and T. Wiederhold

UHMUG-15 *The UHMFS Computer Software*
G. W. Batten, A. Karachievala and H. H. Nguyen

UHMUG-16 *Miscellaneous Computer Software for the Mug File Project*
G. W. Batten

FOREWARD

This report contains a description of several small experiments which are related to the mug file problem. These experiments were not part of the formal research plan for the Mug File Project, but they provide some information which is of interest to other researchers and potential users of our system. For this reason I have encouraged the authors to prepare this report, and I appreciate their efforts.

It should be recognized that these experiments were done as "extra assignments" and did not receive the time and resources allocated to the major tasks of our study. Even though the data is limited, we feel the information content is useful and should be reported.

Ben T. Rhodes, Jr.
Project Director

ABSTRACT

Four experiments were carried out investigating a number of factors in a facial recognition task. These experiments were intended to supplement the work done to develop a computer-based criminal identification system.

Experiment 1 shows that changing three types of accessories (beards, hair length, glasses) on photographs has a strong effect on recognition. This implies that a witness may not be able to identify an out-of-date mug shot or a suspect with a skillful disguise.

Experiment 2 considers target and witness sex, with male-male, male-female, female-male, and female-female combinations. No important differences were found.

Experiment 3 provided the most exciting and unexpected result. In this experiment subjects who had worked as witnesses in an earlier image generation experiment were called back 6-12 months later to see if they could pick out the photograph of their target from a set of photographs. The 19 witnesses made virtually perfect scores in the recognition of the target whom they had seen only once for a 7-10 minute period. This suggests a "stamping in" process has resulted from helping a sketch artist or technician generate an image of the target.

In Experiment 4 subjects used an Identi-kit image or a sketch drawing generated by a unknown witness in an attempt to pick a photograph of the target froma set of photos. The data indicate this is a difficult task, but it is certainly possible to "narrow the list" of possible suspects on this basis.

ACKNOWLEDGMENTS

Many people have contributed to the work reported here. Our colleagues at the University of Houston, Ben Rhodes, George Batten, and Jim Bargainer have provided many ideas in our discussions. Glen Duval has been a constant companion and rich source of ideas and wisdom throughout these efforts. Tony Marchese, an undergraduate student helped carry out the experiment on target and witness sex, and Ron McGowen and Jay Tombaugh, also undergraduate students, helped collect data for the experiment using images as stimuli.

To these and others we express our gratitude.

Kenneth R. Laughery
Richard H. Fowler

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

This document contains a report of a research effort that dealt with facial recognition. The research was part of a larger project to develop a man-computer interactive system for criminal identification. The overall system is described in a separate project report, UHMUG-1.

An important factor in criminal identification concerns the memory that a witness has of a target person. A standard procedure in one type of identification is to have a witness search through a set of photographs, a mug file, attempting to find a match for a face in his/her memory. The task is basically a recognition procedure, where the witness is looking at pictures of faces and making decisions about whether or not each face is the target person.

Previous research has examined a number of factors that may or may not influence facial recognition in a task similar to the mug file search. Laughery, Alexander, and Lane (1971) studied such variables as: the number of facial photographs seen before encountering the target picture; the pose position (front, profile, or portrait) of the faces in the search series; and color versus black and white photographs. In another series of experiments, Laughery, Fessler, Lenorovitz, and Yoblick (1974) explored the effects of elapsed time between seeing the target and beginning the photograph search as well as those of the similarity between the target face and the decoy faces that precede the target in the search series. From these various experiments, two variables emerge as having noteworthy effects

upon recognition performance. The number of photographs seen before encountering the target picture had a considerable impact--the more faces seen, the poorer the performance. The decoy-target similarity also had a strong effect with performance decreasing as similarity increased. The other variables had a relatively minor effect on performance including, somewhat surprisingly, the time lapse between seeing the target and beginning the search.

Another variable that has received considerable attention in facial recognition is the relationship between target race and witness race (Ellis, 1975). Although the results tend to be complex and involve familiarity and attitudes, in general, recogniton is better within races than across races.

In the context of the man-computer interactive system project, several additional experiments were carried out on facial recognition. Actually, a series of four separate studies explored different aspects of the recognition process. The first experiment dealt with a task variable not previously explored; namely, changes in target accessories between the initial exposure and the subsequent exposure in the search series. Specific accessories manipulated were glasses, beards, and hair style.

The second experiment examined the effects of the sex of the targets and witnesses. Previous work on this variable had been inconclusive, some studies showing women might be slightly better witnesses while other studies showed the opposite. In this experiment both male and female targets were used and both males and females served as witnesses.

The third experiment was actually a follow-up to another study carried out as part of this project. In that study a large number of witness subjects worked with either sketch artists or Identi-kit technicians to generate images of a target person they had previously seen. The target populations included White males, Black males, and White females. Details of the image generation study are described in a separate report, UHMUG-2, of this project. The follow-up experiment reported here involved recalling a number of the witness subjects at a later date to participate in a recognition task where the target face was the person whose image they had worked on earlier.

The fourth and final recognition experiment in this series dealt with a different aspect of the criminal identification process and was also a follow-up to the image generation study. In this experiment subjects were shown an image (sketch or Identi-kit composite) of the target person and then searched through a series of photographs of faces attempting to identify the picture of the target person.

CHAPTER 2

EXPERIMENT 1: EFFECTS OF CHANGING ACCESSORIES

An important set of variables in a criminal identification system concerns differences between the target's appearance in the initial exposure and his appearance in a subsequent recognition task. Such changes are no doubt commonplace in actual mug file searches. If the target's picture in the file is very old, there will be normal age-related changes. There may be differences associated with weight loss or gain. If styles have changed or the target used disguises, one would expect differences.

The study reported here explored one class of changes, differences in accessories. Accessories refer to parts of the face that are not permanent and are 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 subjects 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). Subjects first viewed four sequentially presented slides of the target person in different candid positions. The subjects' task was to indicate, using a six-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 manimpulated 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 combinations of conditions used are given in Table 1.

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

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

TABLE 1

ACCESSORY CONDITIONS

Feature	During Exposure	During Search
Beard	yes	no
	yes	yes
	no	no
	no	yes
Hair	long	long
	long	short
	short	long
	short	short
Glasses	yes	no
	yes	yes
	no	no
	no	yes

Only one accessory was manipulated at a time, with the accessory conditions used by four different targets.

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. Then 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.

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 one 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

Figure 1

Sample Target With Different Accessory Conditions

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 subjects in each of the 48 experimental conditions were run as a group. Five subjects 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 subjects. 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 subjects 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 subjects 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. The instructions are presented in Exhibit 1 of the Appendix. 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 subjects recorded their responses on answer sheets.

Any subject who knew the target person was given credit for participation and excused from the experiment. The subjects were asked to indicate on their answer sheets if they knew any of the decoys. There was a negligible number of responses indicating any subject 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 subject the hit-miss (H-M) score could be a single value from 1 to 6. A score of 6 indicates that the subject 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 subject. 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 2. 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 subjects who had a hit (marked a 4, 5 or 6 when the target appeared). These percentages are shown in parentheses 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 idiosyncrasies 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 subject'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

TABLE 2

MEAN HIT-MISS SCORES AND PERCENT OF SUBJECTS

WHO HAD A HIT (IN PARENTHESES)

	Unchanged		Changed	
Accessory	With-with	Without-without	With-without	Without-with
Glasses	4.85 (82.5)	5.45 (92.5)	4.08 (65.0)	4.63 (77.5)
Hair Style	5.35 (90.0)	5.30 (87.5)	3.30 (47.5)	4.13 (67.5)
Beard	5.50 (92.5)	5.10 (82.5)	3.28 (50.0)	3.23 (52.5)

TABLE 3

MEAN FALSE ALARM-CORRECT REJECTION (FA-CR) SCORES

	Target with		Target without	
Accessory	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

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 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, jawline, 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 target's 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 subjects 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 subjects 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.

CHAPTER 3

EXPERIMENT 2: EFFECTS OF TARGET AND WITNESS SEX

In an earlier sequence of experiments, Laughery, Alexander, and Lane (1971) used both male and female witness subjects in a facial recognition task where all target faces were White males. The results were inconclusive, giving no clear indication that either sex was likely to be better witnesses.

In the image generation study done as part of this project (Project Report UHMUG-2), witness subjects worked with sketch artists and Identi-Kit technicians to produce images of a target person they had seen earlier. In two separate experiments White male and White female targets were used. When the quality (match to real face) of the images was compared across the populations, the fit for males was better than for females.

From these results one might conclude that men and women are equally good (or poor) witnesses in recognizing a target person and that men are more readily identified than women. Such conclusions would be shaky at this point, however, since the Laughery, Alexander, and Lane (1971) work actually produced conflicting results - one study showing slightly better performance with male witnesses and another showing females slightly better. Also, the image generation task might not necessarily imply that recognition memory for male faces is better than for females. It could be that the image differences were due to the construction techniques or that recall memory is better for males but recognition memory is not.

The present study consisted of two small experiments designed and carried out in an attempt to resolve these questions and issues. To the extent that either men or women are better witnesses or are more readily identified as targets, the results would have implications for criminal identification systems.

Method

The two experiments employed essentially the same task, procedures, and materials used in earlier recognition experiments (Laughery, Alexander, and Lane, 1971, and experiment 1 in this report). Subjects were exposed to a target person and then saw a series of facial photographs (slides) and indicated whether each was or was not the target.

Subjects. All subjects were students in introductory psychology courses at the University of Houston. Participation earned extra credit in the courses. Thirty-three subjects participated in each experiment, 18 males and 15 females in the first, and 14 males and 19 females in the second.

Task and Procedure. In the first experiment subjects viewed a sequence of four slides showing different candid positions of the target. Following this initial exposure, subjects were shown a series of 55 slides of faces, male or female depending upon the sex of the target. For each face in the search series, subjects indicated on a six-point scale whether that person was or was not the target. The target picture appeared once, in position 45 or 53. The subjects were run in four groups with both males and females in each group. All groups performed the task twice, once with a male target and once with a female target. Two separate

male targets and two separate female targets were used. Two groups saw a male target first and two saw a female first. All other aspects of the task, procedure, materials, and apparatus, were essentially the same as experiment 1 in this report. The instructions to subjects are presented in Exhibit 2 of the Appendix.

In the second experiment subjects viewed one slide showing the target in a front bust pose. The subsequent search series consisted of 135 facial photographs with the target appearing once in position 128. Responses to each face were again indicated on a six-point scale. The subjects were run in groups with both males and females in each group. Each group was presented with one target and search series. Eight groups were run with three variables manipulated orthogonally. The three variables, each with two levels, were target sex, target (2 different males and 2 different females), and initial exposure duration (4 seconds or 8 seconds). The primary reason for this second experiment was that the general level of performance in the first study was high and may have masked the effects of the target sex and subject sex variables. The longer search series and shorter initial target exposure were intended to increase task difficulty. Other aspects of the experiment were similar to the studies referenced earlier.

Results

Hit-Miss (H-M) and false alarm-correct rejection (FA-CR) scores were calculated for each subject (see experiment 1 for a description of these measures). The mean H-M scores from the two experiments were presented in Tables 4 and 5. Although there is a tendency in the second experiment for male subjects to perform better than females, none of the target of subject sex effects

TABLE 4

MEAN HIT-MISS (H-M) SCORES FOR FIRST EXPERIMENT ON TARGET SEX AND SUBJECT SEX (NUMBER OF SUBJECTS SHOWN IN PARENTHESES)

		Target Sex	
		Male	Female
Subject Sex	Male	6.0 (18)	5.7 (18)
	Female	5.9 (15)	6.0 (15)

TABLE 5

MEAN HIT-MISS (H-M) SCORES FOR SECOND EXPERIMENT ON TARGET SEX AND SUBJECT SEX (NUMBER OF SUBJECTS SHOWN IN PARENTHESES)

		Target Sex	
		Male	Female
Subject Sex	Male	5.0 (7)	3.9 (7)
	Female	3.5 (11)	3.5 (8)

were statistically significant in either experiment. Similarly, there were no significant effects with the FA-CR measure.

Discussion

The results of these two experiments are consistent with earlier findings in the sense that there is no clear-cut superiority of either sex as subjects in a facial recognition task. Also, the data indicate that the sexes do not differ in terms of how readily they can be identified. It would appear that sex is not a useful factor in estimating the reliability or usefulness of a witness.

CHAPTER 4

EXPERIMENT 3: RECOGNITION MEMORY FOR TARGETS FOLLOWING IMAGE GENERATION

In most previous work on facial recognition, witness subjects are exposed to a target person and then subsequently attempt to identify that person in a set of alternatives. The initial exposure to the target may involve seeing the person live, seeing photographs of the person, or seeing film or video scenes in which the target appears. Following this exposure, either immediately or after some delay, the witness is shown pictures of faces and asked to identify the target.

The present research project presented an opportunity to explore an interesting question regarding the witness subject's experience with the target and the effect of this experience upon subsequent recognition. Specifically, a series of image generation experiments had been carried out in which 262 witness subjects had worked with a sketch artist or Identi-kit technician to generate an image of a target person previously seen. The target subjects included White males, Black males, and White females. The question of interest concerns the effect of having produced the image on later recognition of the target.

This question is potentially of considerable interest in law enforcement procedures. If, for example, the image generation effort amounts to an intensive rehearsal that "stamps in" the face, subsequent recognition could be considerably enhanced. If, on the other hand, the effort to generate the image (which invariably is far from perfect) results in some confusion about the true appearance of the target, later recognition may actually be impaired.

The experiment as carried out must be viewed as pilot or exploratory work since only a limited number of subjects were tested. Also, all witness subjects had generated images and there was no comparison group that had seen the target under similar conditions but not produced an image. There is, however, a great deal of data available from earlier studies where conditions other than the image generation effort were similar. Comparisons with such data would have to be made informally, of course, and any conclusions drawn with appropriate qualifications. Nevertheless, large differences in recognition performance would be noteworthy, and at the very least would argue for a more careful study to confirm the outcome.

Method

Target Exposure. Details of the witness subjects' exposure to the targets are described in Report UHMUG-2 from this project. Basically, the exposure was a 7-10 minute conversational interaction moderated by an experimenter. In no case had the witness seen the target prior to entering the room where the conversation took place, and the witness never saw the target again after leaving the room.

Subjects. Nineteen people participated in this experiment, all having previously served as witnesses in the image generation study. Many of the witnesses from the earlier studies had been students from the University of Houston or volunteers from the Houston community. Records had been maintained which contained their addresses and phone numbers. A large sample was contacted by phone and nineteen agreed to participate. The elapsed time

between the subjects' participation in the image generation study and their return for this experiment ranged from approximately 6 months to 1 year.

Table 6 presents the image generation experiment condition for each of the subjects. Eleven subjects had generated composites and eight had worked on sketches. Regarding target populations, 4, 6, and 9 subjects worked on White males, Black males, and White females, respectively. The average rating of image quality (a goodness-of-fit measure between the image and target picture) is shown for each of the subjects.

Task and procedure. The recognition task and procedures were essentially the same as in experiment 1 (the accessory study) and as in a series of earlier experiments (Laughery, Alexander, and Lane, 1971). The subjects viewed a series of slides each containing a face and indicated whether each one was or was not the target. The same six-point scale was used to record responses. The slide series consisted of either White males, Black males, or White females, depending upon the race and sex of the target. The subjects were run individually.

The instructions read to the subjects are presented in Exhibit 3 of the Appendix.

The search series for the first three subjects consisted of 44 slides with the target in position 39. For the fourth subject the series contained 58 faces with the target in position 48. The search series for the 15 remaining subjects consisted of 75 slides with the target in position 69. The reason for these modifications was that performance was excellent with the first

TABLE 6

IMAGE GENERATION DATA FOR WITNESS SUBJECTS

USED IN RECOGNITION STUDY

(See Key Below Table)

Witness Number	Target Number	Target Race-Sex	Image Type	Image Quality Mean Rating
116	62	WM	Composite	3.71
115	62	WM	Sketch	3.67
118	63	WM	Composite	3.96
56	35	WM	Sketch	4.02
226	113	WF	Composite	4.38
251	125	BM	Sketch	3.38
276	138	BM	Sketch	3.79
257	128	BM	Composite	4.54
202	101	BM	Sketch	--
208	104	WF	Composite	--
217	108	WF	Sketch	3.42
233	116	WF	Sketch	3.33
216	108	WF	Composite	4.71
245	122	WF	Composite	3.71
230	115	WF	Composite	4.42
214	107	WF	Composite	4.88
252	126	BM	Composite	3.83
281	140	BM	Sketch	3.33
200	100	WF	Composite	--

Key

Target Sex-Race

WM - White Male

BM - Black Male

WF - White Female

four subjects--almost no errors. The series was lengthened in order to increase task difficulty so as to get a better indication of how good the memory was.

The decoy (non-target) faces were randomly selected from the faces slide library at the University of Houston.

Results

The results of this experiment were unexpected--they are virtually perfect. Only four false alarms were committed (a "yes" on one of the decoys), two by each of two subjects. Three of the false alarms were rated "possibly yes" and one "probably yes". In all but one case, the target was identified with certainty; the one exception rated the target "probably yes." Similarly, with only a very few exceptions, all decoys were rejected with maximum confidence.

No statistics were performed on the data because it was not necessary.

Discussion

The results were both unexpected and striking. After a 7-10 minute exposure to a target and working on an image of that target, recognition memory for the person was virtually perfect after time periods ranging from six months to one year. In all earlier studies where images were not generated but the recognition task was begun immediately, performance did not approach this perfect level. Obviously, the image generation task has a strong "stamping in" effect on memory.

The implications of this result are important. By having a witness work on an image of a criminal, one may be able to enhance

considerably that witness' ability to identify the criminal at a later time.

CHAPTER 5

EXPERIMENT 4: RECOGNITION MEMORY WITH IMAGE AS A STIMULUS

Law enforcement procedures often include having a witness work with a sketch artist or Identi-kit technician to prepare an image of a criminal. These images may in turn be used to identify suspects. Indeed, the research project of which these experiments are a part includes using such images as a basis for identifying look-a-likes in a mug file.

This experiment was a follow-up to the image generation study referred to earlier. In that study a large number of sketches and Identi-kit composites were obtained from witness subjects. This experiment explores the ability of other subjects to identify a target person on the basis of having seen a sketch or composite image of him. Given that the image generation experiment included ratings as to the quality of the images, it is also possible to examine recognition performance as a function of image quality.

Method

In most respects the procedures in this experiment were similar to those employed in experiments 1, 2, and 3.

Subjects. Eighty students from an introductory psychology course at the University of Houston served as subjects. Participation resulted in extra class credit.

Task and Procedure. As in the earlier experiments, subjects were exposed to a target person after which they viewed a series of facial photographs (slides). For each face in the search series, subjects indicated on a six-point scale whether or not that face was the target. The search series consisted of 75 photographs,

and the target appeared once, in position 69.

The target exposure consisted of looking at a photograph (a print) of a sketch or composite for 20 seconds. A total of 8 different targets were selected so as to represent two levels of image quality--good or poor. The actual target numbers from the image generation study and the corresponding image quality is shown in Table 7.

Subjects were run in groups of five with two groups for each target. Instructions to subjects are presented in Exhibit 4 of the Appendix. Other aspects of the task, procedure, materials, and apparatus are essentially the same as experiments 1, 2, and 3.

Results

Hit-miss (H-M) and false alarm-correct rejection scores were computed for each subject. The mean H-M scores for the four experimental conditions are shown in Table 8. Analyses of variance on the H-M and FA-CR scores showed no significant effects of either image type or image quality.

The recognition performance of the subjects in this study cannot be compared directly with performance in earlier experiments because of differences in experimental conditions. Yet, it is worth noting that the mean scores (3.0-3.9) were in the same range as was reported by Laughery, Alexander, and Lane (1971) where the target exposures consisted of four casual photographs. Also, they are not unlike the scores reported in experiments 1 and 2, especially in the more difficult task conditions.

TABLE 7

TARGET EXPOSURE INFORMATION FOR EXPERIMENT

USING IMAGES AS STIMULI

(IMAGE QUALITY IS MEAN RATING)

Target Number	Image Type	Image Quality
40	Sketch	2.63-Good
88	Sketch	2.38-Good
82	Composite	2.58-Good
80	Composite	2.71-Good
46	Sketch	4.58-Poor
79	Sketch	4.04-Poor
13	Composite	4.17-Poor
81	Composite	4.25-Poor

TABLE 8

MEAN H-M SCORES IN RECOGNITION EXPERIMENT

WITH IMAGES AS STIMULI

Image Quality	Image Type	
	Sketch	Composite
Good	3.9	3.3
Poor	3.0	3.5

Discussion

It is not clear why the image quality variable did not significantly affect recognition performance. One possible, but not very satisfying, explanation is that the facial information used in rating quality was different from the information used in the recognition task. The reason this notion is not satisfying is that we do not know what the information differences might be. An alternative explanation considers the relationship between image quality and the H-M scores. The ratings for the "good" and "poor" images, while representing the best and poorest matches obtained in the image generation study, nevertheless fall somewhere between the extremes of a perfect match to the target and no match to the target. The H-M scores for the conditions of the experiment are all within 1 scale point. The experiment conducted may not have been sufficiently powerful to reliably detect differences between "good" and "poor" images, as suggested by the means being in the right direction.

The absolute performance level indicated that the task is not impossible. One can recognize faces on the basis of sketches or images. This point is important, of course, since it represents a procedure widely used in law enforcement.

CHAPTER 6

GENERAL DISCUSSION

Four experiments were carried out investigating a number of factors in a facial recognition task. Three experiments explored the effects of changing target accessories, target and witness sex, and the fact that subjects had previously assisted in generating an image of the target. The fourth experiment explored recognition in a situation where the initial exposure consisted of seeing a sketch or Identi-kit composite of the target.

The results show clearly that changing accessories, glasses, beards, or hair style has a strong effect on recognition. The implications of this effect are considerable. Certainly the reliability of an identification is reduced when an accessory change has occurred. An interesting possibility for identification procedures is that computerized systems such as those developed in this project should add a capability for introducing or deleting accessories.

As was true in earlier studies, neither target nor witness sex appears to be a factor in facial recognition. As noted in the introduction, an image generation study showed better images for male targets than for female targets. This apparent inconsistency may be due to the fact that the recognition task and the recall procedures used in generating images may rely on different aspects of the witness' memory. Alternatively, the sex difference in images may be due to technique factors. This possibility makes sense when one notes that the sex difference

was greater for the Identi-kit than for sketches. The absence of any witness sex effect, along with the results of earlier research, indicates that women and men do not differ in any meaningful way in their ability to recognize faces.

The fact that subjects were able to perform the recognition task with a moderate level of success after viewing a sketch or composite indicates that it is worthwhile to disseminate images.

Perhaps the most significant and interesting outcome of these experiments was the virtually perfect recognition performance of subjects after working on an image of the target. A more formal, controlled study is currently underway to confirm this result. If it is confirmed, the implications for law enforcement procedures are important. By simply having the witness work on an image, or perhaps some other form of rehearsal, one can greatly enhance that witness' ability to recognize correctly the target person.

REFERENCES

Ellis, H.D., "Recognizing Faces." British Journal of Psychology, 1975, 66, 409-426.

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.

APPENDIX

INSTRUCTIONS TO SUBJECTS

EXHIBIT 1

Instructions to Subjects in Experiment 1

Effects of Changing Accessories on Facial Recognition

(Subjects are seated and answer sheets passed out.)

In this study we are interested in how well people recognize faces. First, you are going to see four different slides of the same person. Then, you will be asked to identify this person in a series of faces. The sheets you have been provided are to record your identification judgments, and I will tell you how to use them after you have seen the person.

Now I'm going to show you four slides of a person. Try to remember what he looks like, since, as I've already noted, you will be trying to identify him later. Please study only his permanent facial characteristics,since changeable items such as clothing or background may not be a reliable clue to identifying him.

(Four casual photographs of target are shown.)

Now let me explain the use of your answer sheets. You will see that there are seven columns; six labelled from "definitely no" to "definitely yes", and one labelled "known". These columns reflect the kind of decision you will be making about each picture. There are 75 rows; each row will correspond to one slide.

For the first slide you will be seeing, decide if it is the same person you just saw. If you are sure it is the same, place a check on the first row in the column labelled "definitely yes". If you think it is the same, but are not sure, place a check mark

in either the "possibly yes" or the "probably yes" column. Similarly, if you are sure it is not the same person, place a check mark in the column labelled "definitely no". If you think it is not the same, but are not sure, place a check in either the "possibly no" or the "probably no" column. If you know the person personally, check the "known" column. Then, the second slide will be shown; you will make the same judgment about that picture, and mark your judgment on the second row; and so on until all 75 slides have been shown.

Each slide will be shown for seven seconds. To be certain you are on the correct row I will call out the slide number about every 25 items. Do not delay too long in marking your answers, as there is a limited time before the next slide is shown.

Do <u>not</u> assume the target person will appear exactly once in the series of photographs... He may appear once, several times, or he may not be in the series at all. Finally, pay no attention to what others are writing. This is not a test of intelligence or aptitude; we are only interested in your ability to recognize faces.

Are there any questions?

(Show search series.)

EXHIBIT 2

Instructions to Subjects in Experiment 2

Target and Witness Sex in Facial Recognition

In this study we are investigating people's ability to recognize human faces. First you will see four different slides of the same person. Then, you will be asked to identify this person in a sequence of slides of faces.

Please look at these four slides. Try to remember what the person looks like, since you will be trying to identify him later. Please study only his facial characteristics, since his clothing may not be a reliable clue to identifying him.

(Show sequence of target slides.)

Now you will see a series of pictures one at a time. Each row on the answer sheet is for a different slide. For each picture, decide if it is the same person you just saw in these four slides. If you are sure it is the same, place a check mark in the space labelled -Definitely Yes-. If you think it is the same, but are not sure, place a check mark in either the -Possibly Yes- or the -Probably Yes- column. Likewise, if you are sure it is not the same person, place a check mark in the column labelled -Definitely No-. If you think it is not the same, but are not sure, place a check mark in either the -Possibly No- or the -Probably No- column. You may ignore the column marked -Known-.

Each slide will be shown for 5 seconds, followed by a blank for 5 seconds, then the next slide, etc. To help you keep in

the proper place, I will call out the slide number every 25 items.

Please do not delay too long in marking your answers. Although speed is not important in this experiment, we are interested in your first impression of each slide. Also, do not assume the target will appear exactly once in the sequence. He may appear once, several times, or not at all. Finally, pay no attention to what others are writing. There are no right or wrong answers, and we are simply interested in your judgment of the pictures, so please do your own work. Are there any questions?

(Show search sequence.)

EXHIBIT 3

Instructions to Subjects in Experiment 3

Recognition Memory for Targets Following Image Generation

Thank you for coming back for this follow-up experiment. It has been some time since you participated in our original experiment. At that time, you were introduced to another person for the first time. This person we called the target. In our experiment, you spent from 7-10 minutes with this person, and then with the help of either a sketch artist or Identi-kit technician you generated an image of the target.

We would like to determine what effect a long period of time has on the ability of people to recognize someone whom they met for only a brief time. To do this we have collected the slide photos of each target as well as some slide photographs who were not targets in our experiment and assembled them in this slide projector. What we want you to do is indicate, as each picture is shown to you, how sure you are that the picture is or is not the target you described. To do this you have been provided with a response sheet which has a number of rows on each side and seven columns per row. At the top of each column there is a short descriptive label which indicates a degree of sureness. The columns are labelled in the following order: definitely no, probably no, possibly no, possibly yes, probably yes, definitely yes, and known. As each slide is projected onto the screen, indicate your degree of sureness for that slide with a check in the appropriate column. If you know the person being shown, indicate so by checking the known column.

The person you saw may be in this sequence once, more than once, or not at all. The slides are presented at a constant pace, so try to keep up with this pace. If you should fall behind, ask me to indicate what row you should be on. If there are no questions, we will begin.

EXHIBIT 4

Instructions to Subjects in Experiment 4

Recognition Task After Seeing Sketch or Composite

In this study we are interested in how well people can recognize a person after having seen an image, a(n) (sketch/ Identi-kit composite), of that person. On your desk is a photograph of a(n) (sketch/Identi-kit composite) of a person you will try to identify later. Do not look at the photograph now. The sheets you have been provided are to record your identification judgments, and I will tell you how to use them after you have looked at the photograph. You will have 20 seconds to study the image. Please study facial characteristics.

You may now begin studying the photograph.

-(20 seconds)

Please place the photograph face down on your desk.

(As remainder of instructions are read, the photographs are collected individually by E.)

Now let me explain the use of your answer sheets. You will see that there are seven columns; six labelled from "definitely no" to "definitely yes", and one labelled "known". We are going to show you a series of photographs (on slides) of faces. The columns on the answer sheet reflect the kind of decision you will be making about each picture. There are 75 rows; each row will correspond to one slide.

For the first slide you will be seeing, decide if the person is the same individual you just saw in the (sketch/composite) photograph. If you are sure it is the same, place a check on

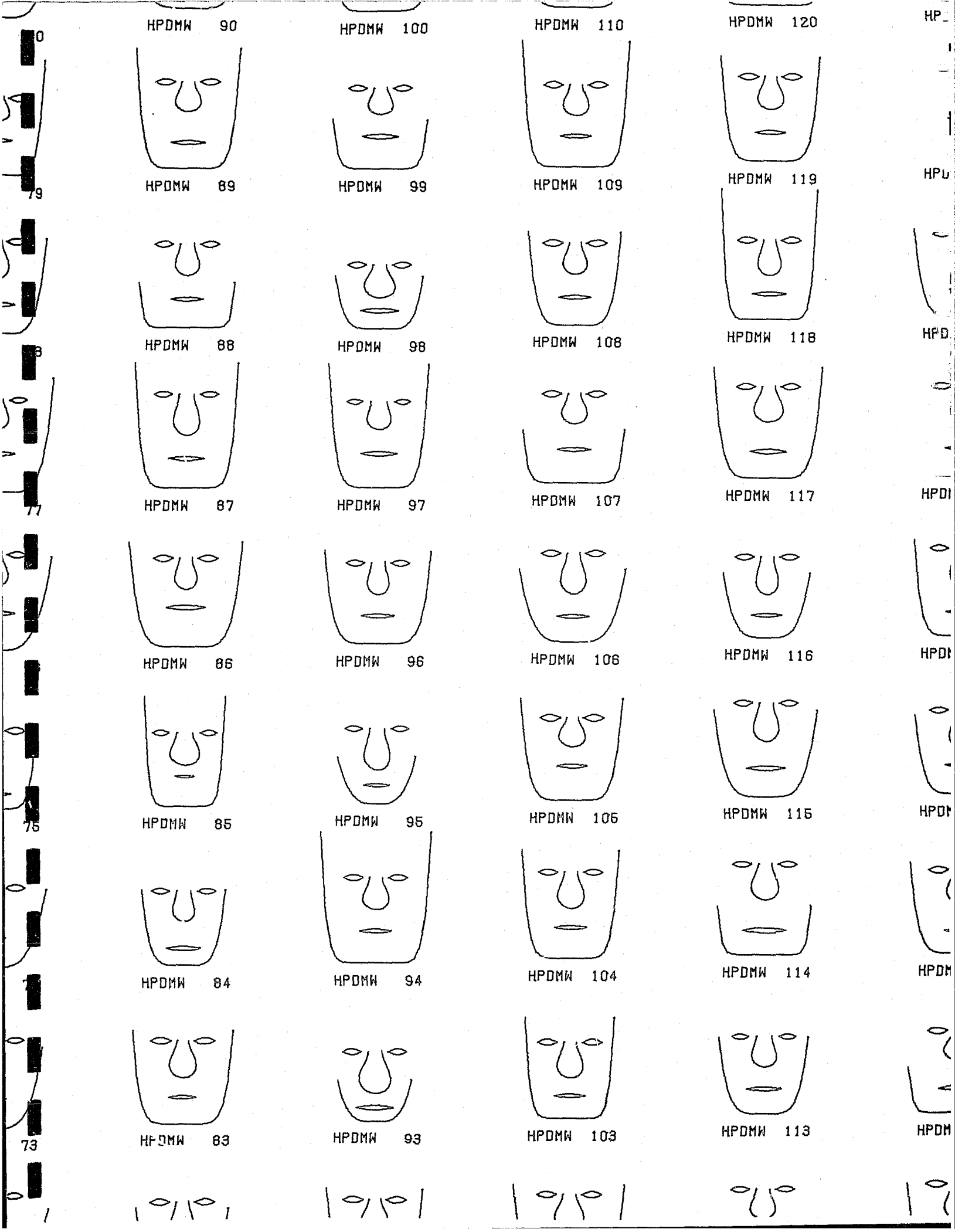
the first row in the column labelled "definitely yes". If you think it is the same, but are not sure, place a check mark in either the "possibly yes" or the "probably yes" column. Similarly, if you are sure it is not the same person, place a check mark in the column labelled "definitely no". If you think it is not the same, but are not sure, place a check mark in either the "possibly no" or the "probably no" column. If you know the person in the slide personally, check the "known" column. Then, the second slide will be shown; you will make the same judgment about that picture, and mark your judgment on the second row; and so on until all 75 slides have been shown.

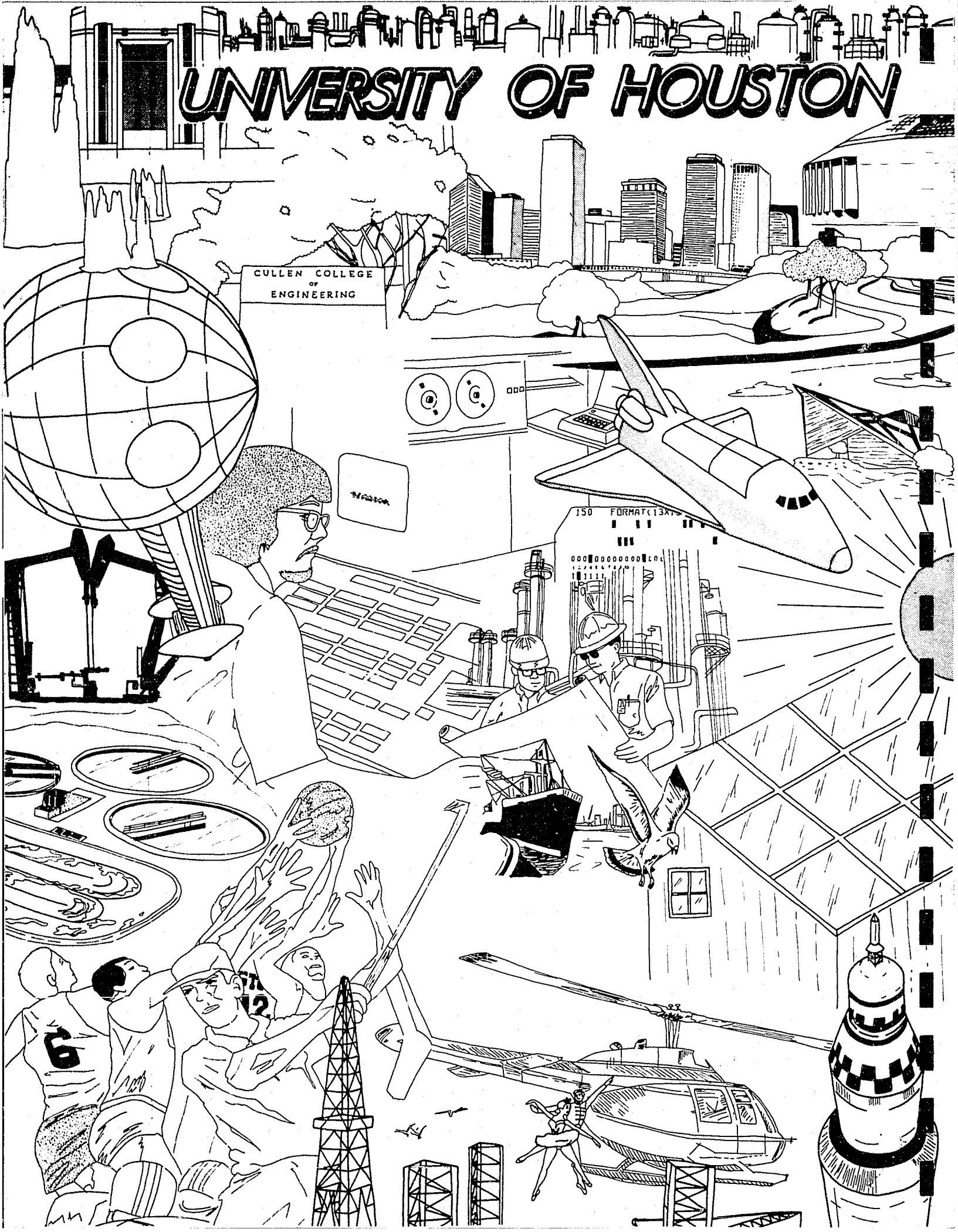
Each slide will be shown for eight seconds. To be certain you are on the correct row, I will call out the slide number about every 10 items. Do not delay too long in marking your answers, as there is a limited time before the next slide is shown.

Do not assume the target person will appear exactly once in the series of pgotographs... He may appear once, several times, or he may not be in the series at all. Finally, pay no attention to what others are writing. This is not a test of intelligence or aptitude; we are only interested in your ability to recognize faces.

Are there any questions?

(Show search series.)


HPDMW 90
HPDMW 100
HPDMW 110
HPDMW 120
HPDMW 89
HPDMW 99
HPDMW 109
HPDMW 119
HPDMW 88
HPDMW 98
HPDMW 108
HPDMW 118
HPDMW 87
HPDMW 97
HPDMW 107
HPDMW 117
HPDMW 86
HPDMW 96
HPDMW 106
HPDMW 116
HPDMW 85
HPDMW 95
HPDMW 105
HPDMW 115
HPDMW 84
HPDMW 94
HPDMW 104
HPDMW 114
HPDMW 83
HPDMW 93
HPDMW 103
HPDMW 113


UNIVERSITY OF HOUSTON
CULLEN COLLEGE
OF
ENGINEERING
150 FORMAT(13X,
6
12

END