4-1-1951

The relative stimulus value of voices and faces as measured by the method of recognition

William Hill Doub

THE RELATIVE STIMULUS VALUE OF VOICES AND FACES AS MEASURED BY THE METHOD OF RECOGNITION

BY

WILLIAM HILL DOUB, JR.

A THESIS
SUBMITTED TO THE GRADUATE FACULTY
OF THE UNIVERSITY OF RICHMOND
IN CANDIDACY
FOR THE DEGREE OF
MASTER OF ARTS IN PSYCHOLOGY

JUNE, 1951

LIBRARY
UNIVERSITY OF RICHMOND
VIRGINIA

TABLE OF CONTENTS

PREFACE

It is often true that initial studies in a little investigated field develop from a broad concept into a more limited area of inquiry. Many begin with the recognition of the need for further knowledge; they strive to focus attention on the problem, and to point the way to further research.

This thesis follows a similar pattern. Although some specific conclusions may be made from the experiments conducted, its chief value is as the initial study in a series - as a beginning rather than a conclusion. Developing from an interest in facial recognition in general, it became in the experimental situation a comparison of the stimulus values of unfamiliar faces and voices.

Grateful acknowledgement is made of the cooperation of the members of the psychology classes who

served as subjects for the pilot study and experiments, and also to the employees of State-Planters Bank and Trust Company who recorded their voices for use in the first experiment.

Much appreciation is expressed to Dr. Merton E. Carver, head of the Department of Psychology, to Mr. Austin E. Grigg, and to Dr. Stanley Skiff, not only for their assistance and cooperation in the preparation of the thesis, but also for their inspiration and encouragement during the undergraduate and graduate studies.

May, 1951 WHD,Jr.

I

INTRODUCTION AND BACKGROUND

When Ebbinghaus published the results of his experimental investigations of memory in 1885, his introductory chapter pointed out the "bare knowledge of the existence of memory and its effects."[1] Historically, in studies of memory our information came in large measure from the observation of the extreme and especially striking cases. The difficulties of scientific studies and the indefinite unspecialized knowledge of the nature of memory was almost prohibitive in undertaking new investigations.

Since the time of Ebbinghaus, however, many hundreds of studies have been conducted in almost

---

[1] Hermann Ebbinghaus, Memory, A Contribution to Experimental Psychology (1885), trans. H. A. Rueger & C. Bussenius (New York: Columbia University Press, 1913), p. 3.

every facet of the phenomenon of memory. Most of the more careful psychological work, especially that which has an experimental basis, has been concerned with special problems in the general field; for example, with the normal course of learning and forgetting, with the influence of special conditions such as position in a series, intensity of stimulation and the like, with classifications of the typical kinds of association and with a study of 'association strengths'.[2]

It is this sort of problem with which this thesis is concerned --- a special problem in the general field. The writer's interest lies in an area where little investigation has occurred, but in which there is promise of practical application. While many studies have been conducted using several variables of auditory and visual memory, relatively few have been directly concerned with the memory for voices and faces.

Individual differences in people's ability to recognize voices and faces is quite popularly accepted. The possession of this ability in some degree is so constantly used and assumed that it is often perceived only

---

[2]Frederick C. Bartlett, Remembering: A Study in Experimental and Social Psychology (Cambridge University Press, 1932) p. 186.

when there is a marked lack of it. A sense of personal identity of people for one another is one of the fundamental bases of social orientation. It is generally accepted that though names and incidents fade and features change, a sense of having seen a face before remains. It is very difficult to analyse the cues used in this type of recognition since, subjectively at least, a general impression is used. For this reason and because of the difficulty of measuring individual factors concerned it is hard to be scientific in studies of the ability. Nevertheless, it is this writer's purpose to analyse the memory for voices and for faces in an experimental situation, and to see what, if any interrelationships are involved. A more detailed statement of the problem will follow, but first a glance at other studies in this general area will serve to make this problem more meaningful.

Again referring to Ebbinghaus, we find his statement that in contrast to the bare knowledge we have of the nature and effects of memory there is an "abundant knowledge concerning the conditions upon which depend the vitality of that inner survival as well as the fidelity and promptness of the reproduction."[3] He points out such

---

[3]Ebbinghaus, op. cit., p. 4.

things as the tremendous individual differences --- not only from person to person but within the same individual when different phases of existence are compared; the great influence of differences of content of the thing to be remembered; the intensity of attention and interest which was attached to the situation the first time it was present; the frequency of repetition; etc. All these have been subjects of inquiry.

In a survey of studies in the specific area of visual and vocal memory, Carlson and Carr[4] as well as Kraweic[5] review numerous experiments comparing memory efficiency of the different senses --- vision, audition, kinaesthesis, and so on. One of the more prolific areas has been the comparison of the effectiveness of visual and auditory modes of presentation on learning and memory. In 1894, Bingham and Munsterberg conducted their classic studies from which they concluded that visual presentation was superior to auditory for learning nonsense syllables.[6] Whitehead, in

---

[4] H. B. Carlson & H. A. Carr, "Visual and Vocal Recognition Memory", Journal of Experimental Psychology, XXIII (1938), 523-530.

[5] T. S. Kraweic, "Comparison of Learning and Retention of Materials Presented Visually and Auditorially," Journal of General Psychology, XXXIV (1946), 179-195.

[6] H. Munsterberg & J. Bingham, "Memory," Psychological Review (1894)I, 34-38, as quoted in Kraweic, op.cit.

similar studies, also reported visual presentation superior to auditory for learning but found auditory superior for retention.[7] On the other hand, Henmon, also working with nonsense syllables, found auditory presentation best in his original studies,[8] but O'Brien in repeating them found no reliable difference.[9] Cantril and Allport favor visual presentation for retention of more difficult verbal material as measured by recall;[10] McDougall also found recognition and recall scores slightly favoring the visual presentation.[11] There seems to be not general agreement as to which modality is best for learning, recognition, or recall of verbal materials. Koch enumerated some of the possible causes of contradictory findings to be

> (1) the different measures of learning efficiency used, (2) the stage in the learning process at

---

[7] L.G.Whitehead, "A Study of Visual and Aural Memory," Psychological Review, III (1896), 258-269, as quoted in Kraweic, op.cit.

[8] V.A.G.Henmon, "Modes of Presentation," Psychological Review, XIX (1912), 79-96, as quoted in Kraweic, op. cit.

[9] F.J.O'Brien, "Qualitative Investigation of the Effect of Mode of Presentation on the Process of Learning," American Journal of Psychology, XXXII (1921), 249-283, as quoted in Kraweic, op.cit.

[10] H.Cantril & G.W.Allport, The Psychology of Radio.(New York: Harper, 1935), p. 165.

[11] R. McDougall, "Recognition & Recall," Journal of Philosophy, Psychology & Scientific Method, I, (1904) 229-233, as quoted in Kraweic, op.cit.

> which measurement is made, (3) the nature of the recording response and (4) the degree of familiarity with the experimental methods used. It is also possible that vision is relatively superios for some individuals and vocality for others and that the two cancel each other for the group as a whole.[12]

While Koch has pointed out these possibilities as explanations for the contradictory findings, Carver suggests that some experimenters have failed to recognize many variables which affect the situation. In the light of this he conducted more comprehensive studies of the roles of four variable conditions. In summary his conclusions were that

> (1) The effectiveness of auditory presentation tends to vary inversely with the difficulty of of the material presented. (2) The effectiveness of auditory presentation is limited to meaningful material, and tends to be superior for subject matter that is concrete and serial in nature. (3)If other conditions are kept constant, the mental functions of recognition, verbatim recall, and suggestibility seem more effectively aroused in listening; whereas critical attitudes and discriminative comprehension are favored by reading. The human relationship involved in the auditory situation is of value for certain types of communication where the personal factor customarily plays a role. (4) The higher the cultural level of the listener the greater is his ability to profit from auditory presentation.[13]

Leaving the more general field of auditory and visual memory and coming to the specific area of this

---

[12]H.L.Koch, "Some Factors Affecting the Relative Efficiency of Certain Modes of Presenting Material for Memorizing," American Journal of Psychology, XL (1930), p. 376.

[13]Cantril & Allport, op.cit., p. 159.

problem, a survey of the literature emphasized the paucity of experiments involving these abilities. Only three were found which were rather closely connected or which followed similar procedures. These are described in detail in order to clarify procedures and to indicate some conclusions already reached concerning memory for voices and memory for faces.

McGehee [14] reports a study on the reliability of the identification of a human voice conducted in connection with the case of the State vs. Hamptmann in 1937. A reader behind a screen read a 56 word passage to a group of auditors. After varying time intervals the original reader and four others read the same selection and the auditors were asked to identify the original reader. This study was later continued except that recorded voices were used rather than live ones.[15] Five men's voices (chosen for similarity of regional speech habits and absence of peculiar dialect or noticeable speech defects) read the same 56 word passage. After varying time intervals again the auditors were asked to identify the original voice. The results of the two

---

[14] F. McGehee, "Reliability of Identification of the Human Voice", Journal of General Psychology, XVII (1937), 249—271.

[15] F. McGehee, "An Experimental Study of Voice Recognition," Journal of General Psychology, XXXI (1944), 53-65.

studies are shown in Table I.

TABLE I: RELIABILITY OF IDENTIFICATION OF HUMAN VOICE

Interval	% Correct Identifications	
	Actual Voice	Recorded Voice
2 days	83%	85%
2 weeks	68%	48%
1 month	57%	47%
2 months	46%	45%

An extension of the second study was an effort to discover why one voice is remembered better than another and whether imagery is used in remembering voices. The conclusions from this section might be summarized as follows: (1) that there was no general agreement as to the most unique or most agreeable, (2) that there was perfect agreement on the lowest pitch but divided on the highest pitch; (3) that there was agreement on the slowest voice but division on the fastest; (4) that the most agreeable voice at first was not generally rated most pleasant after five repetitions; (5) that although imagery was definitely used it was inaccurate in almost every case...that is, there was general agreement as to age, weight, height, personality, and vocation of the stimulus voice, but the judgements were incorrect.

Another study significant in the area of visual memory is one reported by Bartlett.[16] This experiment was quite similar to the present problem in material used but it employed a very different procedure. Five picture post cards, each showing the head of a naval or military officer or man, were presented to the subjects singly for 10 seconds with instructions to note as many characteristics as possible for a later description and questioning about them. After the presentation and the passing of a 30 minute interval (of conversation or other work), the subjects were requested to describe the cards in the order presented and to answer questions concerning some of the details. The recall period was repeated after a week, and two weeks. Of the twenty subjects, seven gave an incorrect sequence at the first recall session, but only one made further errors. Most of these errors were made by the subjects who used visual imagery alone without employing association of the names assigned to the various pictures. It is interesting to note that 60% of the reports of direction of regard were in error at the first recall session. Affective attitudes appeared to color the descriptions considerably and to produce rather stereo-

---

[16]Bartlett, op. cit., pp. 47-59.

typed and conventional reproductions. The experiment was conducted during World War I when there was widespread interest in the armed forces and when fixed conceptions of "type" in the service were rather generally held. A great deal of transference of detail from earlier to later cards was noted in the majority of the cases. Bartlett concluded from this study that in this type of recall for faces, accurate recall is the exception and not the rule.

The third experiment which bears strong similarity to the latter section of this problem is a study conducted by Howells at the University of Colorado in 1938 of the ability to recognize faces.[17] He used six photographs of each of 42 people comprised of 28 women and 14 men, with an age range of 20 to 58. Three different poses of each person were mounted on cards which were presented individually to the subjects. The other three views were presented individually to the subjects. The other three views were presented on one chart divided into 42 rectangles of the same size as the original cards. After seeing each card for ten seconds, the subjects were required

---

[17] T. J. Howells, "A Study of Ability to Recognize Faces," Journal of Abnormal Psychology, XXXIII (1938), 124-7.

to select the same individual from the groups on the chart. This method of presentation necessitated the use of general impressions to recognize the individual rather than specific common elements as cues. Howells found a reliability of .88 for 134 subjects using this procedure. Although the differences were not significant, some indication was found that women were superior to men, and that fraternity students performed better than non-fraternity students. Salespersons' scores were significantly superior to students. The scores correlated .27 with intelligence, .33 with grades, .24 with Allport A-S scales, and .14 with the perception of geometric forms. It was also noted that masking the lower half of the face lowered the scores more than masking the upper half. This confirms Dunlap's statement that "eyes are not the most distinguishing feature of personality contrary to popular notion."[18] The subjects were asked to list the cues which they used for recognition. Those who used the overall impression were generally superior to those who could remember and name more isolated details.

The methods used in the studies cited indicate that a wide variety of procedures is possible in the investi-

---

[18] K. Dunlap, "Role of Eye-Muscles and Mouth-Muscles in the Expression of Emotion," Genetic Psychological Monographs, II (1927) 197-233, as quoted in Howells, op.cit.

gation of the phenomena of memory for voices and faces. There is no one measure of learning.

> There are many ways of measuring retention --- by active recall, recognition, reconstruction and relearning--- and no one measure gives a purer or truer picture of memory than the others... Each of the scores is valid in respect to the particular type of performance for which it stands.[19]

This statement of Postman suggests that the experimenter has several possible procedures available, with recognition and recall being the two most outstanding ones historically. Upon a closer examination of the nature of recognition, recall, and their relationships, it becomes apparent, however, that the nature of the problem to be investigated and the stimuli used will dictate the procedure to be employed in most cases.

Edgell in her discussion of theories of memory defines memory as "cognition of something known before."[20] Preceiving, recognizing, recalling are all psychological functions which belong to the same general series. One recognizes that not everything perceived is necessarily recognized or recalled. From this

---

[19] Leo Postman, "An Experimental Comparison of Active Recall & Recognition," American Journal of Psychology, LXI (1948)511.

[20] Edgell, Theories of Memory.(Oxford: Clarendon Press, 1924),p.145.

> we might reasonably suspect that the differentia of recognition and recall are given, at least partially , in the mode or conditions of the prior perception.[21]

Recognition of things occurs in all degrees of complexity from bare familiarity with a totality to a conceptual analysis of likeness and differences. Bergson considers recognition to be where perception and pure memory are interlaced. Recognition means knowing what is perceived and does not necessarily involve the representation of past experience.[22]

It is generally accepted fact that remembering or recall is a more complex process than recognizing. For instance, in an experimental series only a small portion that can be recognized can generally be recalled. Words and memory images play more prominent parts in recall than they do in recognizing where the immediate stimuli is connected with the sensory pattern.[23] Material recalled usually has to be set in relation with other material, and must be dated, placed, and given some kind of personal mark---some associative value.

---

[22]Bartlett, op. cit., p. 188.

[23]Bergson, H., Matter and Memory (1911) as quoted in Edgell, op.cit.

[24]"The essential difference between recognizing and remembering lies, however, not in an increase of complexity of the latter, but in a genuine difference in the way in which the necessary setting or scheme comes into play." Bartlett, op.cit., p. 195.

The very nature of this problem - i.e., memory for voices and memory for faces - necessitates the use of recognition as the method of measuring this special phase of memory ability. With voices, such a tremendous language difficulty would arise with any use of recall as to make such use impractical. Where memory for faces is involved, the writer's interest is in the stimulus value as measured by recognition, not recall, which would involve associative processes.

Cognizant of this background, it is now possible to state the problem of this thesis:

1. to investigate the impression value for unfamiliar voices;
2. to investigate the impression value for unfamiliar faces presented a) individually and b) in a group;
3. to determine whether or not there is any advantage for voices or faces as stimuli from the viewpoint of impression value; and
4. to determine the amount of measurable retention after one presentation of the two abilities immediately and after a time lapse.

II

PROCEDURE

Before a concrete set of experiments to test these abilities could be divised, certain major questions of the kinds of stimuli and procedure to employ had to be resolved. The primary factors for consideration were 1) whether recordings of voices could be used, 2) whether photographs could be used and, if so, what kind, and 3) how best to present the stimuli in order to get the maximum amount of useful data. It was desirable to conduct the experiments so that they would, in so far as possible, closely resemble real life conditions rather than an artificial laboratory situation.

A major question affecting the recognition of unfamiliar voices was whether the recorded voice was equivalent to the actual voice. In a review of the literature, McGehee's[25] study of voice recognition, conducted at the University of Illinois, provided an answer to this problem.

---

[25] supra, p. 7.

The results of this experiment were compared with an earlier identical study using the actual human voice rather than the recorded voice, and very little difference was found. The author states

> Results of the present investigation indicate that recorded voices may be used in making a psychological study of voice recognition since the difference in recognition of actual and recorded voices amounts to only 7.3% less accuracy over a period of two months.[26]

This agrees with a study by Cantril and Allport in which they found only 7% less accuracy after two months when the recorded voice was used.[27] Use of the recorded voice also has the very important advantage of being identical every time it is heard whereas the actual voice varies with each repetition in spite of all efforts to keep it at a constant speed, pitch, and intonation.

Because of these facts it seemed practical and advisable to use a wire recorder in presenting the voices for this experiment.

Many of the same factors which influenced the choice of the recorded voice over the actual suggested that

---

[26] McGehee, op. cit., p. 55.

[27] Cantril & Allport, op. cit., pp. 109-181, passim.

photographs of faces be used instead of live subjects. Since the problem was recognition of faces any use of live subjects would introduce so many variables as to render accurate controls beyond the realm of probability. Such cues as mannerisms, expression, movement, etc., would nnnecessarily complicate the experimental situation.

The photographs selected were from annuals of three large colleges. By using pictures of seniors of 1944 or earlier, all of whom were non-residents of Richmond, the chances of prior familiarity of the stimuli to the subjects were sbbstantially eliminated. All the pictures were 1½ x 2 inches and were presented to the subjects by projection on a screen so that they appeared slightly larger than life-size. Only pictures with clear focus were chosen so that projection would not cause distortion.

With these major questions resolved, attention was given to other details in preparing the experimental material.

The stimulus voices used were employees of a local bank who were white adults between the ages of 18 and 30. They were selected with attention to several factors: there were no exceedingly deviant accents althgugh there was

a range fromslightly northern diction to the mild Southern drawl; no noticeable speech defects were used; they spoke clearly and distinctly; and they represented the same type of people found at the university where the experiments were conducted.

The selection of passages to be read posed several problems. It was desirable both from the point of view of wording and length to have a maximum of cues in a minumum of time. Thus a passage was chosen in which the words were familiar and the thought a complete one. It also offered some chance for expression in reading. Because of the extreme difficulty of equating passages, it was considered mandatory to use one for all the stimulus voices and one for all the test voices. In order to reduce fatigue from repetition of the same passage, a humourous incident was chosen. The final selections were short fillers, 20 to 25 seconds reading time, from Coronet magazine:

Stimulus voices

> A young lady stepped into a drugstore and asked how to take castor oil without tasting it. The druggist said he'd look up some suggestions, but meanwhile, would the young lady relish a refreshing lemonade? She would. When the beverage was entirely consumed he asked laughingly, "Well, did you taste it?"
>
> "Good heavens !" gasped the girl. "Was the castor oil in that lemonade? I wanted it for my mother."

Test voices

> Mother and daughter were very busy with the wedding plans when the bridggroom-to-be called. He watched the preparations rather impatiently until his future wife noticed his look of annoyance.
>
> "Darling, we have such a lot to do," she soothed, "and if we want to make our wedding a big success we mustn't forget even the most insignificant detail."
>
> "Oh, don't worry about that," murmured the young man. "I'll be there all right."

Very brief instructions were given to all the readers:

> Read this passage through several times to familiarize yourself with it and then read it in your normal tone into the mike as if you had found it amusing and were reading it to a friend.

In selecting the photographs for the experiment involving recognition of faces presented individually, factors were considered similar to those which influenced the voice stimuli selections. Of the individual photographs, no persons who had particularly striking features or outstanding feature details were selected. The overall size of the pictures was equal as was the head size. The head and shoulders was shown of each individual. Since it was desirable to eliminate all stimuli other than the faces, cues from position, clothing, lighting, coloring, and direction of regard were kept to a minimum. The manner of dress was held constant in order to reduce cues from this factor; all women wore white blouses of the same design with no jewelry; the men were attired in conservative suits, white shirts, and plain ties. The

photographs were in clear focus with little shadowing and were black and white. The directions of regard were divided almost evenly between 3/4 left and 3/4 right.

These factors also affected the selection of the group pictures. In addition all groups were composed of from 9 to 11 men or women (five all women and five all men groups). No person appeared in more than one picture. All individuals were in clear focus.

Mimeographed forms were devised for the recording of responses. These forms were simply constructed so that the response could be indicated by placing a check or an X mark in the space provided. Samples of these report forms appear in Appendix A. A credit of 1 was allowed for each critical stimulus correctly recognized, making the total number of correct responses the score for the experiment.

The selection of materials completed, a pilot study was run using 130 summer school psychology students as subjects. Three experiments --- I, a test of recognition of unfamiliar voices; II, a test of recognition of unfamiliar faces presented singly, and III, a test of recognition of unfamiliar faces presented in group pictures--- were conducted with an immediate response group and a delayed

response group after a 48 hour time lapse. Upon analysis of the results and some preliminary statistical treatment, several changes in the procedure of the experiments were deemed necessary.

There were no apparent weaknesses in Experiment I, so the same procedure was retained for the main study. Several of the individual stimulus pictures in Experiment II were easily identified by every subject. Subjects reported a distinctive feature or feature detail as the cue to identification. These pictures were eliminated and others substituted which more closely resembled the majority of the test pictures.

The frequency distribution curve for Experiment II evidenced some negative skewness. In order to obtain more normal distribution the time of exposure of the stimulus was cut from 20 seconds to 5 seconds; this was effective and a wider range of scores was obtained in subsequent trials.

Excessive difficulty of Experiment III was indicated by the pilot study. The group poses were analyzed for clearness and slightly longer time limits were allowed. The mean of the initial Experiment III subjects was 1.3 and the highest score was 4 out of a possible 10. After some of

the group pictures were changed and the time of exposure increased, the mean of an equivalent group of subjects was found to be 4.16.

When these changes had been made in material and procedure, the experiments were conducted. First the subjects were given initial instructions; the ten stimulus voices or pictures were presented; then instructions for recording responses were given, followed by the test of recognition.

The subjects used in the main study were students of six general psychology sections of Richmond and Westhampton colleges. With few exceptions they were college sophomores, the average age being 19. There were 171 men and 96 women. Of the total of 267 students participating in the experiments, 136 were in the immediate group while the delayed section numbered 131. Since, however, not all of the subjects were present on all the days the experiments were conducted, the total usable sample included 214 subjects.

There was no indication of practice effects being present in the pilot study. As a safety measure, however, to balance any which might have appeared, the experiments were given in the several combinations: one section of the immediate group and one of the delayed took experiment I, II,

and III; one section of each had II, III, and I, and one of each had III, I, and II.

No further controls than those previously indicated were included. There are perhaps many other factors which affect how these abilities operate, but it is the writer's opinion after consideration of many similar studies that these variables should become the subject of individual studies.[28]

Recognition of Unfamiliar Voices

A wire recorder was used to present ten voices -- five men and five women. These stimulus voices were recorded alternating a man and a woman. In random order the ten stimulus voices were then mixed with ten additional

---

[28]"He who considers the complicated processes of the higher mental life...will in general be inclined to deny the possibility of keeping constant the conditions for psychological experimentation...Factors which are to the highest degree determinative and to the same extent changeable, such as mental vigor, interest in the subject, concentration of attention, changes in the course of thought -- all these are either not at all under our control or are so only to an unsatisfactory extent. However, care must be taken not to ascribe too much weight to these views, correct in themselves, when dealing with fields other than those of the processes by the observation of which these views were obtained."

Ebbinghaus, op. cit., p. 11.

voices (five men and five women) for the recognition section of the experiment.

The following instructions were given to all the subjects at the beginning of the experiment:

> You are going to hear a series of ten voices. They will all say the same thing. Don't bother about the content, just listen to the voice so you will be able to recognize it later. Here is a sample of what you will hear:

The immediate response group received the preliminary instructions, heard the voices, were given instructions in marking their responses and then heard the test recording. The delayed response group received the second instructions and the test of recognition after a 48 hour lapse of time. These second instructions were:

> Now you will hear the same ten voices mixed in with ten others; they will all read the same thing. You are to try to recognize the ten voices you have heard previously. On the sheet of paper you were given you are to place a check beside #1 if you have heard it before. If you are certain, put the check in the column under the word "certain". If you are guessing place the check in the column under the word "guess." If you have <u>not</u> heard the voice before, put an X mark. For example: ...(sample voice)...You heard this voice as the sample before so put a check in the column marked "certain" beside the word "sample."

At the beginning of the experimental record, mimeographed report forms were furnished the subjects and the instructions played. These were followed by a sample voice in order to eliminate the possibility of laughter obscuring the first voice. The test recording was then played

and the subjects completed the record form. At this point, they were also asked to indicate what cues they found most valuable in recognizing the voices they had heard previously.

Recognition of Single Individuals

The second experiment was conducted very similarly to the voice study. Ten pictures of individuals, alternating a man and a woman, were presented as stimuli to the subjects for an interval of 5 seconds per picture. This was followed by the presentation of a series of 20 faces arranged in random order and composed of the ten critical pictures and 10 additional ones. The subjects were asked to select the 10 previously seen.

Again the two groups of subjects - immediate and delayed response - were used. The instructions for both groups were identical with the latter group receiving the test of recognition after a 48 hour time lapse. The instructions were:

> You are going to see a series of 10 pictures. There will be five men and five women. You will see them for about five seconds each. Look at each one carefully so that you will be able to recognize the person later. Here is a sample picture. You will now see the ten test photographs. Look at them all carefully.
> (Presentation of pictures)
> Now you will see these same ten mixed in with ten other photographs which you haven't seen before. You are to try to recognize the ten people you saw previously. On the sheet of paper you were given you are to place a

> check by #1 if you have seen it before. If you are certain, place a check beside #1 in the column under the word "certain". If you aren't sure, but think you have seen it before, place the check in the column under the word "guess". If you have not seen the picture before place an X mark on the sheet. Here is a sample picture...You saw this picture as the sample before, so put a check beside the word "sample" in the column under the word "certain." Now you will see the 20 test photographs. You have seen ten of them before. Put a check by the numbers of the ones you have seen.

After completing the record blanks, the subjects were asked to indicate what cues they had used in recognition of the critical stimulus pictures.

Recognition of Individuals in a Group

Experiment III involved the use of ten stimulus pictures, five men and five women, presented singly as stimuli and followed by the test of recognition using ten group pictures - five of all men and five of all women. The stimulus pictures were exposed for 20 seconds each. The task was to select from each of the group poses individuals who had been presented singly as stimuli. On the completion of presentation of the individual faces, the test of recognition was conducted. The following instructions were used:

> You are going to see a series of ten pictures, five men and five women. You are to look at each one very carefully so that you will be able to recognize the people later. After all 10 pictures are presented you will be asked to pick out the people you have seen before.
>
> (Presentation of stimulus faces)

Now you will be shown ten group pictures. The ten people you saw previously will be in these groups. There may be none in some groups and one or two in others, so look at them very carefully. Each person in the groups is numbered. If you have seen one of the people of group #1, write the number which is on that person, in the space beside #1 on the mimeographed sheet you were given. If you are certain you have seen the person put the number in the column under the word "certain." If you have not seen any of the people in the group, put an X mark. Here is a sample picture of a group - none of whom have you seen before. It is only to illustrate how to record your answer. If you were certain that you had seen the person numbered 4, for example, you would write the number "4" in the column under the word "certain." If you thought you had seen that person, but weren't sure, you would write the number "4" in the column under the word "guess." If you were certain that you had not seen any of the people before, you would put an X mark. After all ten pictures are shown you should have ten numbers on your sheet of paper. Please DO NOT put more or less than ten numbers.

The time lapse of 48 hours was again used for the delayed response group, and all subjects were requested to indicate cues to recognition.

III

PRESENTATION OF DATA

Following the collection of the data through the series of experiments conducted, both logical and statistical analysis was made of the results. Distribution tables were calculated and frequency curves drawn. Data for the immediate and delayed recognition groups was treated separately for each experiment.

Critical ratios were applied to determine if the differences of the means of the immediate and delayed groups were significant. In order to ascertain the consistency of the relationships between the experiments, again the method of critical difference ratio was applied.

The frequency distribution tables for each experimental group will be found in Appendix B. Figure I represents the distribution curves of the data on recognition

of unfamiliar voices. In the immediate recognition group of 125 subjects, the mean score was 5.816. The range of scores was from 3 through 10 and the standard deviation was 1.2421 giving the curve a slightly leptokurtic appearance. This same condition was true for the delayed group where the range of scores was from 3 through 9 and the standard deviation was 1.240. For the 100 subjects of the delayed response group the mean score was 5.96.

FIGURE I: FREQUENCY DISTRIBUTION
EXPERIMENT I --- RECOGNITION OF VOICES

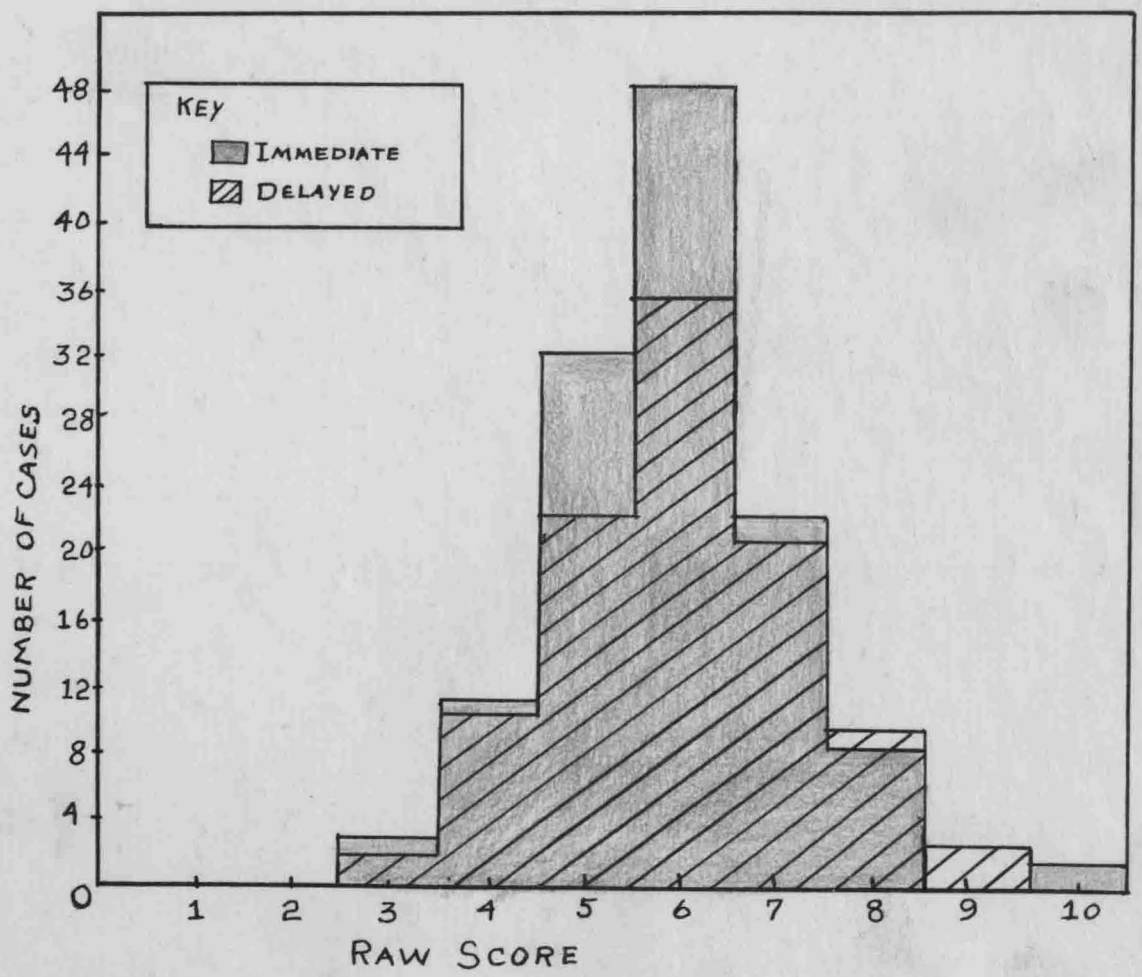
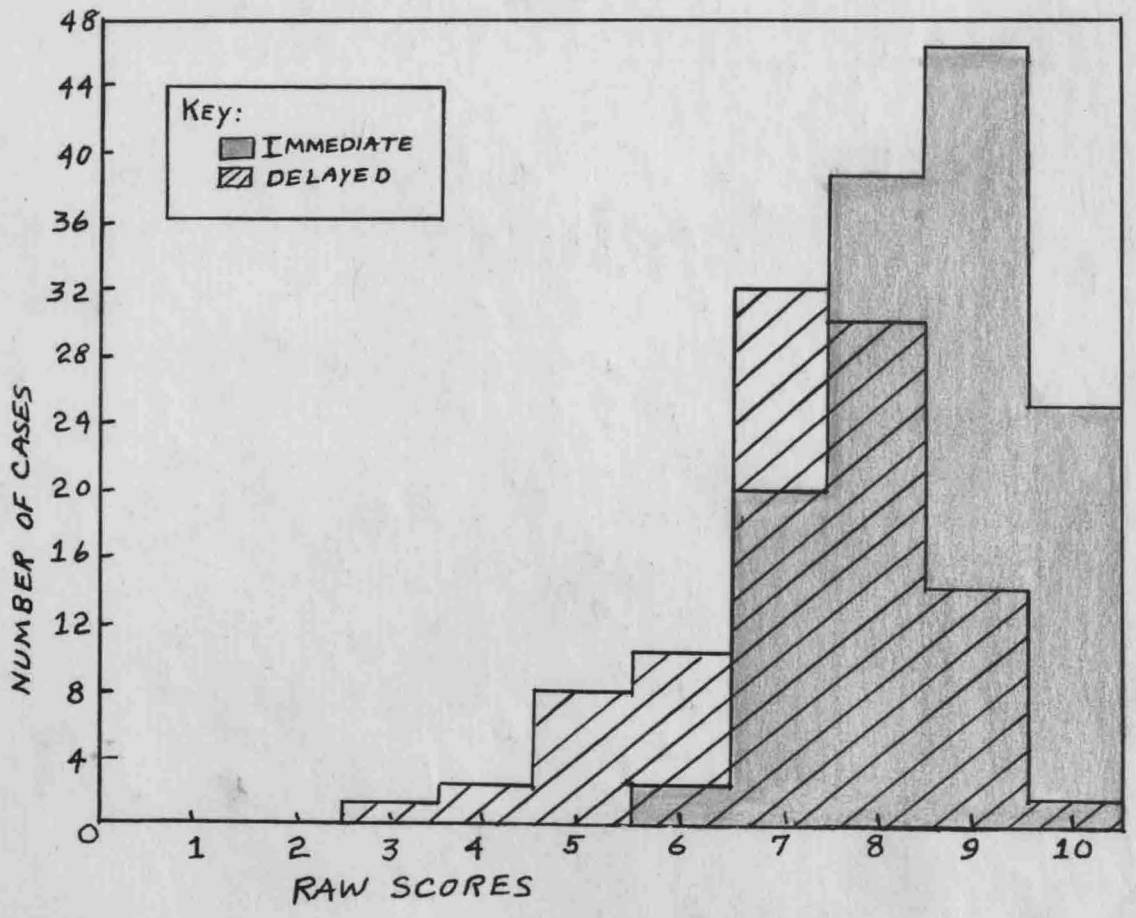


Figure II represents the frequency distribution curves of the data from Experiment II, recognition of unfamiliar faces presented singly in series. The 130 subjects in the immediate response group had a mean score of 8.561. There was a restricted range of scores at the upper end of the scale (from 6 through 10), giving the curve an almost negatively skewed appearance. The standard deviation was 1.012.

FIGURE II: FREQUENCY DISTRIBUTION
EXPERIMENT II --- RECOGNITION OF FACES (single)



The delayed response group had a more normal appearing curve with a mean score of 7.264. The standard deviation of the distribution was 1.273 and the 102 subjects had scores ranging from 3 through 10. From the pilot study it was expected that the scores would pile up nearer the upper limits for the immediate group. Therefore the exposure time of the stimulus was cut to yield a more normal distribution; the addition of perhaps two more faces would have improved the distribution further.

FIGURE III: FREQUENCY DISTRIBUTION CURVE
EXPERIMENT III --- RECOGNITION OF FACES (group)

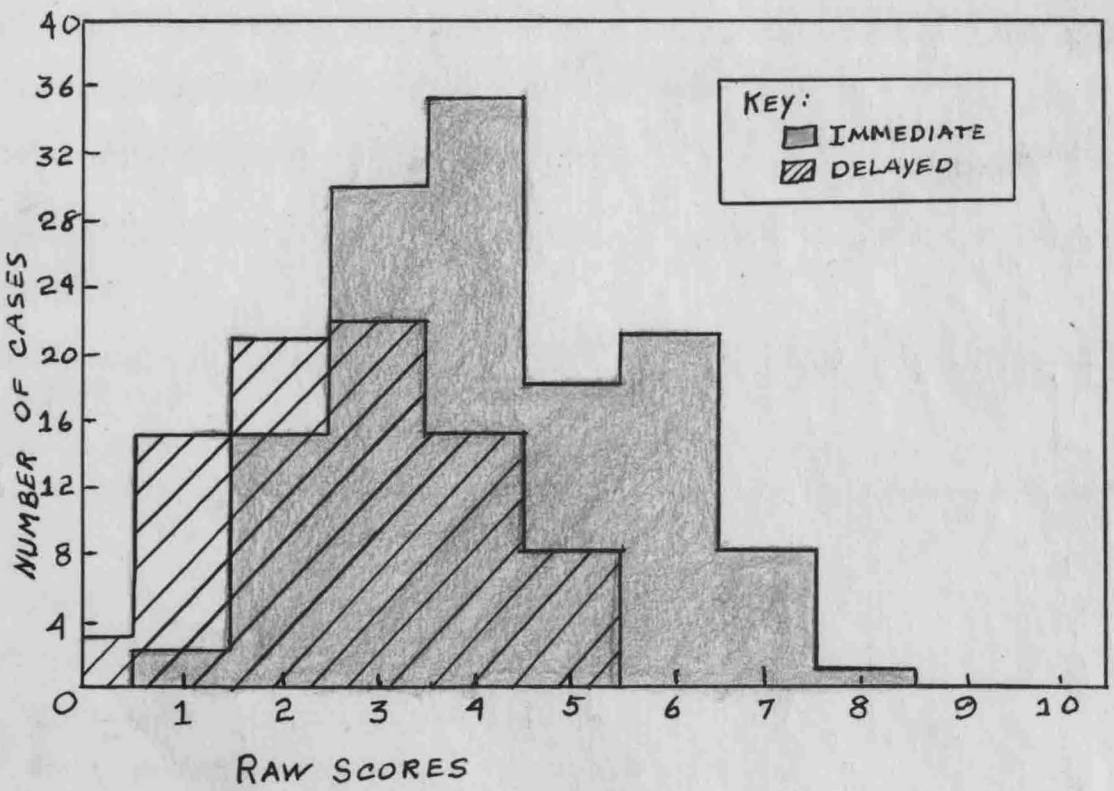


Figure III represents the distribution of the data on recognition of unfamiliar faces presented in a group. The 130 subjects of the immediate response group had a mean score of 4.149. The range of scores was the widest for any of the experiments, varying from 1 through 8; there was a standard deviation of 1.505. This curve approximates a normal distribution better than any of the other experimental groups. For the 64 delayed response subjects, the mean score was 2.654, with scores varying from 0 through 5 and a standard deviation of 1.313.

In order to determine if the differences in the means of the delayed and immediate response groups were true differences, critical ratios were computed[29] by dividing the obtained difference by the standard error of the difference using the following formula.

These critical difference ratios for each experiment appear in Table II.

TABLE II: CRITICAL RATIOS OF IMMEDIATE & DELAYED RESPONSES
(Note: Exp. I - Voices; Exp. II - Faces; Exp. III - Faces(group)

Experiment Number	Mean Immediate	Mean Delayed	Difference of Means	Critical Ratio
I	5.816	5.960	.144	.008
II	8.561	7.264	1.297	.797
III	4.149	2.634	1.515	.758

[29]For sample calculation, see Appendix C.

Chance factors for the three experiments were empirically determined to be 5, 5, and 1, respectively. It is evident from comparing the mean score with the chance score that very little learning took place in Experiment I, voices. It is interesting to note that the mean delayed recognition score was superior very slightly to the mean immediate score. The critical ratio (.008), however, indicates that this was pure chance and is not a reliable difference.

Even though many theories could be advanced for the small amount of learning, the most practical would appear to be that the passages were so short as to afford few cues. This, coupled with the fact that the subjects' ability to discriminate between the slight variations in pitch, tone, rate of reading, expression, etc., would seem to account for the small amount of learning which occurred. It is also probable that auditory cues of this type are more difficult to assimilate and structure than visual cues. It would be interesting to test the auditory discrimination threshold of the subjects with the highest and lowest scores and determine the correlations of auditory discrimination abilities and the ability for recognition of voices.

In Experiment II, recognition of faces presented individually, a larger difference in the means and chance score indicate that more learning took place. It

should be noted that when the same time was allowed as in Experiment I (see pilot study) the mean score for immediate recognition was over 9. Thus it would seem that, for the average subject, the task of recognizing unfamiliar faces is considerably easier. Although this is not a significant difference, it does indicate a trend (79 chances in 100).

Comparing the mean and chance scores in Experiment III, recognition for faces presented in groups, one infers that a greater amount of learning took place than in either of the other experiments. There are several possible reasons. There was a longer exposure of the stimulus picture than in Experiment II; this allowed a more careful study of the stimulus and absorption of the cues. It is perhaps a safe supposition that visual cues are more easily perceived than auditory ones, since it has been shown that 85% of all impressions are received visually. In Experiment III, the differences in the immediate and delayed responses are not significant although the trend (77 chances in 100) is toward a true difference.

To summarize these observations, it may be said that the stimulus value of unfamiliar faces seen once, for the same stimulus period is greater than the stimulus value of unfamiliar voices. In the pilot study experiments with unfamiliar faces presented singly, the mean score of

the immediate response group was above 9 whereas the mean score of the voice experiment was 5.816. When the presentation time for the visual stimulus was cut from 20 to 5 seconds the mean score was still 2.745 above that for recognition of unfamiliar voices. The conclusion can be made therefore that the stimulus value of a human face seen once is greater than that of a human voice heard once when such impressions are measured by recognition.

Analysis of the indicated cues to identification employed by the subjects was very interesting. Most of the subjects suggested generalities as the cues and the majority seemed to use a sense of general familiarity as the basis for recognition. It is of interest to note that in Howells' study of recognition of faces[30] subjects who used the feeling of general familiarity were generally superior to those who selected specific ones. An analysis of the cues used and the score made by the subjects indicates more confirmation than denial of this idea.

The specific cues most frequently mentioned for the voice experiment were accent, expression, rate, pitch, and smoothness of reading. Most of these were mentioned in connection with a particular stimulus voice and could not be employed consistently for every voice. Of course, the

---

[30] supra, p. 11.

initial selection of the stimuli was carefully slanted at eliminating widely deviant voices so that the nature of the stimuli influenced the primary role of a sense of general familiarity in the recognition process.

With the two experiments involving recognition of faces, again the sense of general familiarity (e.g. "seemed familiar", "thought I'd seen it before") was the most frequently mentioned cue. Many subjects, however, also mentioned distinctive features which applied to one or another of the pictures but not necessarily to all. Among the specific cues indicated were eyes, hair, mouth, facial structure and other feature details.

This preliminary logical inspection of cues involved in the memory ability for voices and faces suggests that a further investigation of this problem would be enlightening, for it is the writer's inference that the cues for visual recognition are better organized than those for auditory recognition. Facial stimuli can be structured more easily than vocal ones since the cues are presented simultaneously and in structured form rather than singly and unstructured. This would allow easier visualization and organization of the visual stimuli.

Certainly it is recognised that many people make a conscious effort to remember faces whereas recognition of voices is not a social or professional demand, except for selected groups of people (e.g. telephone operators). The sub-

jects appeared to experience less difficulty in their analyses of cues for faces than those for voices.

Another sidelight of the problem which merits some attention is the relationship of the responses indicated as "certain" and those marked as "guess".

The papers generally fell into these three patterns: 1) those which had a nearly equal division of "certain" and "guess"; 2) those which were nearly all "certain"; and 3) those which were nearly all "guess". There was no relation between the degree of certainty and efficiency of performance. It is the writer's belief that this might be a variable not of the ability of recognition of faces or of voices, but rather of the personality of the individual --- whether he is secure and confident in his judgement. Howells[31] related the recognition of faces to scores on the Allport Ascendancy-Submission Scales and obtained a correlation of .24. The writer believes that a correlation of the pattern of certain-guess responses would be more positively related to the A-S scores than would the actual ability of recognition itself.

Having noted the operation of the separate phenomena in the experimental situations, an analysis of any

---

31 supra, p. 11.

existing interrelationships logically follows.

It will be recalled that the mean scores for recogntition of unfamiliar faces presented singly were superior to the means for recognition of unfamiliar voices which in turn were superior to the means for recognition of unfamiliar faces presented in groups.

In order to find whether the differences in the mean scores between the experiments are reliable differences, critical ratios were calculated[32] and appear in Table III.

TABLE III: CRITICAL RATIOS OF INTERRELATIONSHIPS OF EXPERIMENTS
(Note: Experiment I, Voices; Experiment II, Faces (single); Experiment III, Faces (group)

	Experiment Numbers	Means	Difference of Means	Critical Ratios
Immediate	I & II	M_1 5.816 M_2 8.561	2.745	1.882
	I & III	M_1 5.816 M_3 4.149	1.647	.831
	II & III	M_2 8.561 M_3 4.149	4.932	2.178
Delayed	I & II	M_1 5.960 M_2 7.264	1.304	.720
	I & III	M_1 5.960 M_3 2.654	3.306	1.890
	II & III	M_2 7.264 M_3 2.654	4.610	2.556

[32] For sample calculation, see Appendix C.

It will be noted that although none of the differences are completely reliable, all but voices and faces in groups (immediate) and voices and faces singly (delayed) have better than 96 chances in 100 of being true differences. It is significant that the highest reliabilities were obtained in comparing the experiments on identification of faces.

It was the original intention of the writer to compute coefficients of correlation in order to determine the nature of relationships existing between the experiments. Upon analysis of the data, however, it was felt that critical ratios were more revealing.

Since no learning occurred in the experiment involving voice recognition, the means remained close to chance whereas in the experiments concerning facial recognition the means were significantly above chance. Therefore no correlation could be expected. In order to verify this hypothesis the coefficients of correlation were computed[33] and were found to be: voices and single faces, .003; voices and faces in group pose, .049; and single faces and faces in group pose, .040.

---

[33]For sample calculation, see Appendix C.

IV

SUMMARY AND CONCLUSIONS

Whenever a new area of investigation is selected it is generally true that the results of first experiments take the form of verified observations and limited conclusions rather than any sweeping generalized conclusions covering an entire field. This was true in the areas covered by this problem. Several commonly assumed ideas are verified in the data gathered and certain popularly held ideas are in some measure contradicted.

The writer feels that, although some new light was shed on the problem, the chief value of these experiments has been to focus attention on a problem which has been little investigated but has possibilities in practical applications and to suggest further areas of inquiry and experimentation.

Historically, many studies of memory, its nature, and its effects have been conducted. A large measure of them have dealt with memory efficiency of the various senses. In the field related to the theme of this problem, much of the work has been comparison of memory efficiency for visual and auditory presentation of verbal material, e.g., nonsense syllables, etc. Relatively few of the studies have investigated the problem of recognition of unfamiliar voices and faces through the use of a single presentation procedure and minimal learning.

This field has very definite possibilities of application in testing for aptitudes in meeting and having successful public relationships. Perhaps the most similar application has been the use of memory for names and faces in such tests as Hunt's Social Intelligence Test and the studies of validations of this in various public contact jobs.

This thesis was concerned, however, not so much with the idea of recall and associative processes nor with the relative effectiveness of visual and auditory presentation as such but rather with the relative impression value of faces and voices under experimental conditions that approach in some measure the everyday life situation.

Such studies as these might be used in the construction and validation of tests for such positions as telephone operator, receptionist, wholesale and retail sales personnel, and public relations representatives.

It is obvious that more extensive study is necessary before specific application may be made, but the results of this thesis have suggested certain variables which would be of value in further investigation for the nature of the phenomena and the conditions under which they operate.

Further refinements and extensions of the present study would be of value. Some extensions in which the author's interest has been aroused are 1) conducting a similar study using groups differing in age, educational level, occupation, intelligence, social intelligence, and extroversion-introversion ratio; 2) using motion pictures as well as still pictures in order to approach more closely lifelike situations; and 3) measuring the amount of superiority of stimulus value of faces over voices.

Studies of this sort would be valuable from both a theoretical and a practical point of view. On the basis of this study some conclusions of practical value may be made.

The more important of these is that the stimulus value of a human face seen once is greater than that of a human voice heard once when such impressions are measured by recognition. This could be partially due to the fact that in interpersonal relationships the facial expression is studied constantly in order to pick up cues to the individual's feelings; another contributing factor would be that visual stimuli are more easily structured and organized. Whatever the cause, the findings support the conclusion that the face has a higher stimulus value. This finds practical application in the idea that effort toward recognition may more efficiently be concentrated on learning a person's appearance rather than his voice, since the experiment shows that the visual task is easier and because everyday situations allow for greater practice of the visual task. It may also be noted that visual stimuli may not be.

In summary, that data from these studies support the following conclusions:

1. The stimulus value of the human face seen once is greater than that of a human voice heard once when such impressions are measured by recognition.

2. Very little relationship exists in this experiment between recognition of voices and recognition of faces presented either singly or in groups.

3. No relation was discernable between the degree of certainty of response and efficiency of performance.

4. The majority of the subjects used a feeling of general familiarity as the cue to recognition rather than specific cues.

BIBLIOGRAPHY

Arkin, H. & Colton, R. R., Outline of Statistical Methods, 4th edition. New York: Barnes and Noble, 1932.

Bartlett, Frederick C., Remembering: A Study in Experimental and Social Psychology. Cambridge University Press, 1932.

Cantril, H. & Allport, G. W., The Psychology of Radio. New York: Harper, 1935.

Carlson, H. B. & Carr, H. A, "Visual and Vocal Recognition Memory," Journal of Experimental Psychology, XXIII (1938), 523-530.

Edgell, Beatrice, Theories of Memory. Oxford: Clarendon Press, 1924.

Ebbinghaus, Hermann, Memory: A Contribution to Experimental Psychology (1885). trans. Rueger, H. A. & Bussenius, C. New York: Columbia University Press, 1913.

Garrett, Henry E., Statistics in Psychology and Education. New York: Longmans, Green and Co., 1926.

Howells, Thomas H., "A Study of Ability to Recognize Faces," Journal of Abnormal Psychology, XXXIII (1938), 124-127.

Koch, H.L., "Some Factors Affecting the Relative Efficiency of Certain Modes of Presenting Material for Memorizing," American Journal of Psychology, XL (1930), 370-388.

Kraweic, T. S., "Comparison of Learning and Retention of Materials Presented Visually and Auditorially," Journal of General Psychology, XXXIV (1946), 179-195.

Lindquist, E. F., A First Course in Statistics. Cambridge: Riverside Press, 1942.

McGehee, Frances, "Reliability of Identification of the Human Voice," Journal of General Psychology, XVII (1937), 249-271.

McGehee, Frances, "An Experimental Study of Voice Recognition," Journal of General Psychology, XXXI (1944), 53-65.

Postman, Leo, "An Experimental Comparison of Active Recall and Recognition," American Journal of Psychology, LXI (1948) 511-519.

VITA

WILLIAM HILL DOUB, JR. Born, Richmond, Virginia, February 17, 1929. Educated, Richmond Public Schools, Richmond, Virginia. Graduate, John Marshall High School, Richmond, Virginia, February 1946. Graduate, Richmond College, University of Richmond, Virginia, June 1949, B.S. degree in chemistry. Entered, Graduate School, University of Richmond, September 1949. Appointed, Psychologist, Virginia State Penitentiary, April, 1951.

APPENDIX A

SAMPLE OF REPORT FORMS AND PICTURES

I. Report Form of Experiment I, Voices

Voice	Certain	Guess
Sample		
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		
16		
17		
18		
19		
20		

NAME _____

_____

AGE _____ SEX _____

CUES USED:

II. Report Form for Experiment II, Individual Faces

Picture	Certain	Guess
Sample		
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		
16		
17		
18		
19		
20		

NAME _____

_____

AGE _____ SEX _____

CUES USED:

III. Report Form for Experiment III, Faces, Group

NAME_____

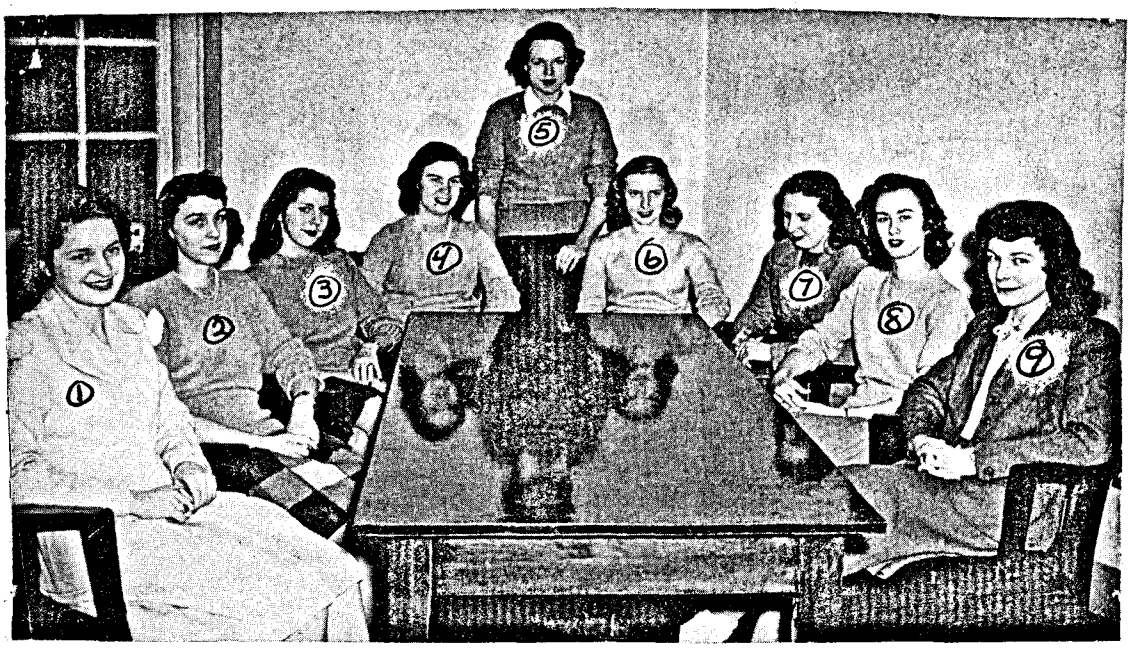
SEX _____ AGE_____

Picture Sample	Certain	Guess
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		

CUES USED:

IV. Sample of Individual Stimulus Photograph - Experiment II and Experiment III.

V. Sample of Group Photograph - Experiment III



APPENDIX B

FREQUENCY TABLES FOR GRAPHS OF FREQUENCY DISTRIBUTIONS.

TABLE I. FREQUENCY DISTRIBUTIONS FOR EXPERIMENT I - VOICES

Raw Scores	IMMEDIATE Number of Cases	IMMEDIATE % of Sample	DELAYED Number of Cases	DELAYED % of Sample
10	1	.8 %	0	.0
9	0	.0	2	2.0
8	8	6.4	9	9.0
7	22	17.6	20	20.0
6	48	38.4	35	35.0
5	32	25.6	22	22.0
4	11	8.8	10	10.0
3	3	2.4	2	2.0
2	0	.0	0	.0
1	0	.0	0	.0
0	0	.0	0	.0
	125	100.0 %	100	100.0 %

MEAN	5.816, Immediate	-	5.960, Delayed
MEDIAN	6.000		6.000
MODE	6.000		6.000

TABLE II. FREQUENCY DISTRIBUTIONS FOR EXPERIMENT II - FACES (single)

Raw Scores	IMMEDIATE Number of Cases	IMMEDIATE % of Sample	DELAYED Number of Cases	DELAYED % of Sample
10	25	19.2	1	.98
9	46	35.3	14	13.7
8	38	29.2	32	31.7
7	19	14.6	34	33.3
6	2	1.5	10	9.8
5	0	0	8	7.8
4	0	0	2	1.9
3	0	0	1	.98
2	0	0	0	0
1	0	0	0	0
	130	99.98	102	100.02
MEAN	8.561		7.264	
MEDIAN	9		7	
MODE	9		7	

TABLE III. FREQUENCY DISTRIBUTIONS FOR EXPERIMENT III - FACES (group)

Raw Scores	IMMEDIATE Number of Cases	IMMEDIATE % of Sample	DELAYED Number of Cases	DELAYED % of Sample
10	0	0	0	0
9	0	0	0	0
8	1	.76	0	0
7	8	6.15	0	0
6	21	16.15	0	0
5	18	13.84	8	9.5
4	35	26.90	15	17.8
3	30	23.07	22	26.2
2	15	11.50	21	25.0
1	2	1.50	15	17.8
0	0	0	3	3.5
	130	99.87	84	99.8
MEAN	4.169		2.654	
MEDIAN	4		3	
MODE	4		3	

APPENDIX C

FORMULAS and SAMPLE CALCULATIONS

I. Standard Deviation

A. Formula:

$$\sigma = \sqrt{\frac{\Sigma(x)^2}{N} - \left(\frac{\Sigma x}{N}\right)^2}$$

where Σx^2 = sum of squares of scores

$(\Sigma x)^2$ = sum of scores squared

N = number of cases

B. Sample Calculation:

$$\sigma = \sqrt{\frac{4421}{125} - \left(\frac{727}{125}\right)^2}$$

$$\sigma = \sqrt{35.368 - 33.826}$$

$$\sigma = \sqrt{1.542}$$

$$\sigma = 1.2421$$

II. Critical Ratio

A. Formula:

$$C.R. = \frac{D}{\sigma \text{ diff.}}$$

where D = difference of means

$$\sigma \text{ diff} = \sqrt{\sigma_{M_1}^2 + \sigma_{M_2}^2}$$

B. Sample Calculation:

$$C.R. = \frac{2.745}{\sqrt{(1.07)^2 + (.99)^2}}$$

$$C.R. = \frac{2.745}{1.458}$$

$$C.R. = 1.8827$$

III. Correlation

A. Formula:

$$r = \sqrt{\frac{[N\Sigma xy - (\Sigma x)(\Sigma y)]^2}{[N\Sigma x^2 - (\Sigma x)^2][N\Sigma y^2 - (\Sigma y)^2]}}$$

where N = number of cases

Σxy = sum of the cross-products

Σx = sum of scores of experiment 1

Σy = sum of scores of experiment 2

Σx^2 = sum of squares of scores of exp.1

Σy^2 = sum of squares of scores of exp. 2

B. Sample calculations:

$$r = \sqrt{\frac{[121(2927) - (102)(503)]^2}{[121(4275) - (102)^2][121(2365) - (503)^2]}}$$

$$r = \sqrt{\frac{(1061)^2}{(14,471)(33,156)}}$$

$$r = \sqrt{.0023}$$

$$r = .049$$