Philos
J

Journal of Experimental Psychology

EDITED BY

Volume IV, 1921

PUBLISHED BI-MONTHLY BY
PSYCHOLOGICAL REVIEW COMPANY
PRINCETON, N. J.

Application for entry as second-class matter at the post-office at Lancaster, Pa., pending

PRESS OF
THE NEW ERA PRINTING COMPANY
LANCASTER, PA.

INDEX FOR 1921

Journal of
Experimental Psychology

Vol. IV, No. 1. February, 1921

THE INSPIRATION-EXPIRATION RATIO DURING TRUTH AND FALSEHOOD

BY HAROLD E. BURTT

Ohio State University

1. Introduction

Three variables have been studied as diagnostic of truth and falsehood: the association reaction, systolic blood pressure, and the inspiration-expiration ratio. Benussi[1] obtained rather clear results with the last of these methods. The writer in collaboration with Marston and Troland made a brief study of all three methods with a view to military use.[2] In this case the breathing results were negative. Since that time a technique has been developed which increases the accuracy and ease of obtaining the I/E (ratio of inspiration time to expiration time). The present study utilized this improved technique to repeat in its essential features and extend Benussi's experiment.

Benussi's subjects were given small cards containing a number of digits and letters and a picture. They were questioned as to the nature, number and arrangement of the symbols and were asked to describe the picture and to read the symbols in a specified order. In some instances they did this truthfully but with other cards (marked with a star) they lied upon every point. A number of spectators were present to enhance the subjects' emotional state. The

[1] Benussi, V., 'Die Atmungsymptome der Luge,' *Archiv f. die gesamte Psychologie*, **31**, 1914, 244–273.

[2] Cf. Shepard, J. F., 'Review of Affective Phenomena,' *Psych. Bull.*, **16**, 1919, 244.

breathing was recorded on a kymograph and several breaths immediately preceding and following each reply were measured to compute the I/E.

Benussi found that the average of 3 to 5 ratios preceding a false answer was less than the average of 3 to 5 ratios following that answer. With truthful replies the reverse was the case,—the average I/E before the answer was greater than the average I/E after the answer. The difference between the average ratios before and after a lie varied with the ability of the subject to dissimulate (as judged by the experimenter) and correlated also with the difficulty of lying in specific instances as indicated by the introspection. The results were negative if the audience knew whether the subject was going to lie or not and if the subject was aware of this. Benussi found that when the subject tried to voluntarily control the breathing, even following a metronome, the I/E was still diagnostic.

2. Method

The subject in the present experiment wore a pneumograph of the usual type,—a large rubber tube with a spiral spring inside. This connected pneumatically with a tambour which closed an electric circuit during inspiration and opened it during expiration.[1] In the earlier part of the experiment this operated a double relay which closed circuits through two signal magnets,—one during inspiration and one during expiration. These two circuits were interrupted by a vibrator 5 times per second. The magnets recorded in a spiral on a kymograph about 12 feet distant from the subject and the I/E was obtained by counting vibrations on the two lines. With a telegraph key in parallel with the relay circuits the experimenter could mark on the record in simple code the number of the question and other data necessary for interpretation. In the later part of the experiment the relay circuits operated a device which indicated the I/E at the close of each expiration[2] so that it could be recorded immedi-

[1] Burtt, H. E., 'A Pneumograph for Inspiration-Expiration Ratios,' *Psychol. Bull.*, 15, 1918, 325–328.

[2] Burtt, H. E., 'Further Technique for Inspiration Expiration Ratios,' Jour. of Exper. Psych. (to appear shortly).

ately. It consisted essentially of two light cars moving at right angles from a common point, one during inspiration and the other during expiration. A pointer was pivoted on the former and rested against a stop on the latter so that at the end of a breath the I/E was a function of the angle covered by the pointer and could be read from a scale. The subject also wore a sphygmomanometer of the portable type and a bracelet stethoscope. Systolic blood pressure readings were taken at intervals.

The experiment consisted of four series. The first employed material quite similar to Benussi's. A set of 3 x 5 inch cards was prepared each containing capital letters or digits or both. There were from 4 to 8 symbols on a card and they were arranged in one of 6 simple geometrical shapes,—triangle, circle, etc. The number, nature and shape of the symbols was determined by chance and 18 cards of the sort were prepared. The subject after 30 seconds study was questioned on four points: the geometrical shape, the kind of material (letters or numbers or both), the number of symbols, and an actual reading of the symbols in some prescribed fashion. The order of these questions was varied. The subject had before him a reference card on which were the six geometrical shapes, a statement of the fact that there were always from 4 to 8 symbols and the four points above mentioned on which he was to be questioned. He was informed that if, when lying, he confined himself to the limits indicated the experimenter could not tell from the answers whether he was lying or telling the truth about a card selected at random from the group. In an hour's experiment eight cards were generally used. Four of these the subject described truthfully and with four he lied on every point. The order was determined by chance and unknown to the experimenter.

The above points were explained to the subject and he was told to try to deceive the experimenter. The latter after each card stated whether he considered the performance was truth or lie. This was primarily to create greater interest and perhaps emotional reaction. One or more spectators were sometimes present for the same purpose. They likewise passed judgment after each card.

In order to obtain breathing records before and after the answers the subject was instructed to remain silent when asked a question, to think about the answer until the signal 'now' after which he could reply as soon as he wished, and to then keep silent until the next reply. The experimenter waited about 10 seconds,—long enough for 3 to 5 breaths,—after asking a question before saying 'now' and waited a corresponding length of time before giving the next question. The subject was told that this procedure was to avoid his betraying himself by the length of time he took in replying to questions. Blood pressure readings were taken at the outset of the hour and after the second and fourth questions on each card.

The method of Series II. was exactly like that of Series I. just described except that the first four cards were all truth or lies and the last four the reverse. The order of the two parts of the hour's experiment was determined by chance and unknown to the experimenter. His judgment had to be deferred, of necessity, until the end of the hour.

In Series III., the subject was given two typewritten sheets of paper, face down, marked respectively T and L. He selected whichever he wished. In either instance the paper gave an account,—comprising at least 10 points on which he could be questioned,—of a crime of which he was supposed to be accused. If he chose the paper marked L the account implicated him in the crime and he had to lie concerning nearly every point in order to clear himself. The paper marked T, however, recounted the same evidence against him but provided an alibi on every point. He had merely to tell this alibi without fabricating. The following is an excerpt from an L and a T paper covering the same 'crime.'

L

You are named as corespondent by Mrs. James Wyandotte in a suit against her husband for divorce. Witnesses have established the following:

1. About Christmas time you began wearing an expensive fur although you had received no raise in salary recently and none of your friends or relatives admitted that they gave it to you.

2. About New Year's you appeared with an expensive necklace.

3. On the evening of January 10 you and Mr. Wyandotte sat in adjacent seats at the Hartman theatre and after that took lunch at the Busy Bee. You were not at home that night.

4. On the afternoon of February 24 part of a telephone conversation which you were carrying on from Hennick's was overheard. About the middle of the conversation you said: "Well what will friend wife say?" . . . etc.

T

You are named as corespondent by Mrs. James Wyandotte in a suit against her husband for divorce.

The truth is as follows:

1. About Christmas you purchased a fur with money which you made through speculation, on the strength of a tip overheard at the office where you work.

2. You purchased a necklace a little later with money from the same source.

3. You went to the Hartman Theatre, January 10, with a girl friend, Miss Helen Reilly, and it was only after you had been seated that you noticed Mr. Wyandotte's presence in the next seat. After the theatre he followed Miss Reilly and yourself to the Busy Bee where he seated himself at the same table with you and tried to flirt with you, and your friend. You spent that night with your friend, Miss Reilly.

4. The telephone conversation which was overheard at Hennick's was with this same girl friend. Your remark about "What will friend wife say?" was in regard to a party which Miss Reilly was going to hold in her rooms. 'Friend-wife' was a name which you called Miss Reilly's landlady. . . . etc.

A considerable variety of crimes were used including murder, theft, arson, forgery, fraud, treason and grounds for divorce.

The L papers were written by the experimenter but the corresponding T papers by an assistant. The experimenter did not examine the T papers and thus could not tell from the subject's replies whether he was making up his own alibi, *i.e.*, lying, or following the one written by the assistant, *i.e.*, telling the truth.

The subject was allowed 5 minutes to study the paper he selected. After each question was asked he waited for the signal 'now' before replying as in the previous series. Blood pressure was likewise measured at the outset and after every three or four questions. Two "crimes" were performed in an hour's experiment. The subject was instructed to select the L in one crime and the T in the other but the order was immaterial and unknown to the experimenter. The latter at the end of the hour made a judgment as before, to insure interest, on the basis of the subject's manner and a casual survey of the blood pressure record.

Series IV. was identical with Series III. in procedure except that a number of 'jurors' (from 3 to 26) sat facing the subject and after each crime judged on the basis of his manner and statements whether it was T or L. This jury did not know that one performance during the hour was to be T and the other L but supposed the subject would do whatever he wished in each instance.

The questioning of the subject in Series III. and IV. was not conducted after the manner of a legal cross examination. A list of questions (unknown to the subject) was prepared before hand and asked in a matter of fact manner. Occasionally a few additional questions were interspersed. Undoubtedly an agressive rigorous examination would produce a heightened emotional state. But the aggressiveness is a variable which depends on the personality of the examiner. The writer was interested in whether this personality variable could be eliminated and still obtain positive results.

In all series for each complete breath between question and answer (there were from 3 to 5 breaths) the I/E was computed and these ratios averaged. Similarly the ratios for the 3 to 5 breaths following the answer were averaged. It seemed best, following Benussi, to average the ratios rather than to average inspirations and divide by the average of the corresponding expirations. (These methods would of course be identical if the geometric mean were to be used throughout but there is no particular justification for doing so.) It was desirable to note the ratio for each separate breath in order to study, for instance, the diagnostic accuracy of earlier and later breaths.

The average I/E for the breaths following an answer was always subtracted from the average I/E for those preceding that answer. For convenience this difference will hereafter be called D. According to Benussi, D should be positive for true answers and negative for false answers. In some cases D was computed, using only one breath before and one after the reply. These cases will be mentioned later.

The experiments were performed in the psychological laboratory of the Ohio State University, during the winter of 1919–20. Three graduate students (2 female, 1 male) served as subjects throughout. A number of other persons served as subjects in one or two series. Of these 2 were graduate students (both male), 7 were undergraduates (2 female, 5 male) and 2 were outsiders (both female).

3. Results of Series I and II

The first two series, it will be recalled, employed cards containing letters and digits similar to Benussi's. Four questions were asked about each card and there were usually 4 L cards and 4 T cards in an hour's experiment, a total of 32 questions. Several aspects of the results of these series are presented in Table I.

Table I

Subject	Av. *D* for 16 Questions		No. of Questions *D* Correct		No. of Cards *D* Correct		B.p. Cards Correct	
	L	T	L	T	L	T	L	T
(Ser. I)								
B.	−.09	.01	8/16	8/16	3/4	3/4	3/4	4/4
B.	.07	.03	3/16	10/16	1/4	2/4	4/4	2/4
C.	.05	−.10	5/11	3/12	0/3	0/3	3/3	3/3
E.	−.11	.09	8/12	9/13	3/4	3/4	3/4	3/4
H.	−.24	−.10	4/4	2/5	2/2	1/2	3/3	3/3
J.	−.18	−.05	8/13	7/15	4/4	3/4	2/4	2/4
K.	−.01	.00	9/14	7/13	2/4	2/4	2/4	2/4
M.	.03	.16	4/10	7/13	1/4	4/4	1/4	2/4
N.	.04	.06	6/16	9/15	1/4	3/4	3/4	3/4
Av. or total	−.05	.01	55/112	62/118	17/33	21/33	24/34	24/34
			49%	53%	52%	64%	71%	71%
(Ser. II)								
B.	.02	.02	7/12	7/11	1/3	2/3	x	
K.	.12	.12	4/13	6/11	1/4	2/4	x	
M.	.07	.05	3/13	9/15	1/4	2/4	x	
N.	.06	−.09	3/14	5/15	1/4	1/4	x	
Av. or total	.07	.03	17/52	27/52	4/15	7/15		
			33%	52%	27%	47%		

The data in each row are based on an hour's experiment with the subject indicated at the left. The *D*'s (average *I*/*E* before question minus average *I*/*E* after question) for the 16 questions in which the subject lied are averaged in the 2d column. Similarly the average *D* for the 16 questions in which the subject told the truth is given in the 3d column. According to the hypothesis *D* should be negative in sign for L and positive for T. Column 4 gives the number of L questions in which *D* is correct according to hypothesis, *i.e.*, negative. This figure is followed by the total number of L questions, because in some instances the record was not complete for all 16 questions. Column 5 gives likewise the number of T questions in which D is correct, *i.e.*, positive, followed by the total number of T questions. These same questions are of course the basis of the averages in the preceding columns. Each card was also evaluated as a unit, *i.e.*, the 4 D's of the 4 questions averaged into a single score. The number of cases in which this score correctly diagnoses an L card are given in column 6 together with the total number of L cards. Column 7 gives similar results for the T cards. The last two columns give the number of cards correctly diagnosed on the basis of systolic blood pressure. These results will be considered separately (section 6).

The upper portion of the table presents the results for Series I. in which the T and L cards occurred in a random order. Columns 2 and 3 are averaged and the remaining columns totaled. With these totals moreover the first figure in each column is taken as a per cent. of the second and these percentages given directly below (for example in the 4th column 55 is 49 per cent. of 112). The lower portion of the table presents the results for Series II. in which the 4 L cards were used in succession followed by the 4 T cards or vice versa. These results are likewise averaged or totalled. With the blood pressure readings it seemed desirable to evaluate not each card but the two groups of cards. The crosses in the last column indicate that the diagnosis as to whether the first or last four cards are L is correct.

The results for Series I. indicate a certain correspondence, although not a striking one, between the I/E and the telling of truth or falsehood. It will be recalled that in this series rather simple and unemotional stimuli were used, the questions were not of the cross-examination type and were known beforehand by the subject. Under these circumstances any correspondence between performance and breathing is encouraging. The second and third columns of Table I. show that for the majority of the subjects D is diagnostic and the average D for the L's is $-.05$ as compared with .01 for the T's. If the average of all the original D's (about 115 of each sort) is used instead of the average of the figures in the table the results are − .06 for L and .04 for T. The difference between these two averages is 4.6 times the probable error of difference and hence significant.

Considering the percentages in the following columns it is evident that in only about half of the original questions (49 per cent. on L and 53 per cent. on T) is *D* diagnostic. This apparent discrepancy with the preceding averages is probably due to the fact that many questions produce little consciousness of deception or truth whereas an occasional one does so to a considerable degree. Many of the original records show a considerable number of *D*'s varying slightly on both sides of zero, intermixed with a few which are of considerable plus or minus magnitude. The percentage of correct diagnoses is slightly greater when considering cards (column 6 and 7) rather than individual questions. It seems plausible that one of the four questions on a card might provoke a significant reaction which would alter the card average.

Two individual points should be noted. Subject *E* performed in the presence of 12 spectators and might be expected to have a heightened lying consciousness, whereas in two other series one spectator was present and in the rest no one but the experimenter. *E*'s breathing appears in the table the most diagnostic of the group. Subject *M* confessed later that instead of fabricating regarding the card before him he kept the preceding card in view and described that for his lie, thus in reality telling the truth. His average shows a positive, *i.e.*, truthful, *D* for both parts of the series.

The relative sensitivity of the first and last questions on each card was studied to determine if the lying consciousness seemed objectively more in evidence early or late during the process. The results are not included in the table because of their doubtful significance. *D* is slightly more diagnostic on the last question but this might well be due to the fact that the last question was usually more complex,—involving a reading of the symbols in some prescribed orders.

The results of series II are less convincing. The averages of columns 2 and 3 show positive values in both cases and even a slightly higher value (.07) for L than for T (.03). The direction of this difference is the reverse of that called for by hypothesis, but a calculation using the original *D*'s shows that this difference between the averages is only 1.8 times the probable error of difference and so of very doubtful significance. The percentages in the following columns likewise show little correspondence between breathing and mendacity. There is frequent occurrence of positive *D*'s in L questions, *i.e.*, breathing indicating truth. (*Cf.* columns 4 and 6.)

Two facts might partially account for the less significant results of Series II. In the first place all the subjects had already done Series I. and perhaps had become more used to the situation and less excited by it. In the second place the four L cards were taken in succession and the lying consciousness might wear off during that process. As a matter of fact the first L card of the 4 shows a negative average *D* for 3 of the 4 subjects whereas the last L card shows a positive average *D* for all subjects.

The results of these two series seem then to indicate that even with simple unemotional material and with questions known beforehand and asked in an unexciting manner there is some evidence of correspondence between the breathing ratio and the telling of the truth or falsehood. This correspondence is generally due to a few questions in the group while the majority of them are neutral. The influence of lying upon the breathing appears to become less with habituation.

4. Results of Series III

In Series III., it will be recalled, the material consisted of imaginary crimes and in a given hour's experiment the subject lied about one crime and told the truth about another. It was thus possible to deal with absolute values or to compare a T and an L performed under practically the same conditions. The results are summarized in Table II.

The data in each row are based on one hour's experiment. The left portion of the line (columns 2 to 7) gives figures for the crime regarding which the subject lied and the right portion (columns 8 to 13) for the crime regarding which he

Table II

Subject	Lie						Truth						Correct Alternative			Correct b.p.
	D_1	No. Cor.	D_2	No. Cor.	D_3	No. Cor.	D_1	No. Cor.	D_2	No. Cor.	D_3	No. Cor.	D_1	D_2	D_3	
K	−.02	5/10	−.08	7/10	.01	5/10	.02	8/12	−.02	7/12	−.03	5/12	x	x	o	x
M	.04	5/11	.24	5/11	.21	4/11	−.01	6/11	.06	5/11	−.07	5/11	o	o	o	x
M	.21	2/8	.16	2/8	.29	1/8	−.13	4/12	.00	7/12	−.02	6/12	o	o	o	x
N	.01	2/10	.06	2/10	.09	1/10	.00	5/11	.06	7/11	.09	7/11	o	?	?	x
N	.02	4/6	−.02	2/6	−.02	3/6	.01	2/4	−.03	0/4	.00	0/4	o	o	x	x
N	−.06	8/12	−.02	5/12	−.03	3/12	−.03	7/12	−.09	2/12	−.11	3/12	x	o	o	o
Q	−.07	4/8	.05	2/8	−.01	3/8	.07	4/5	.06	4/5	.08	4/5	x	x	x	o
Av. or total	.02	30/65 46%	.06	25/65 38%	.08	20/65 31%	−.01	36/67 54%	.01	32/67 48%	−.01	30/67 45%	3/7	2/7	2/7	5/7

D_1: av. of 3 to 5 breaths between question and answer minus av. of 3 to 5 breaths following answer.
D_2: first breath after question minus first breath after answer.
D_3: first breath before answer minus first breath after answer.

told the truth at the same sitting. Symbols representing the crimes are not included in the table because the difficulty and emotional stimulus of the various crimes was an individual matter and impossible to evaluate. The usual procedure is to average 3 to 5 I/E ratios preceding the reply and subtract the average of 3 to 5 ratios following the reply. This is done for each question. The average of these D's on a given crime with a given subject are given in the columns marked $D1$. The immediately following columns (3 and 9) give the number of questions in which this $D1$ is correctly diagnostic (*i.e.*, negative for L and positive for T) together with the total number of questions on which ratios are available. It seemed possible that certain breaths might be more significant than others. Accordingly the I/E for the breath following an answer is subtracted from the I/E for the breath following the experimenter's question. The average of these differences for a given crime are given in the columns marked $D2$. The next columns in each instance (5 and 11) give the number of questions in which $D2$ is correctly diagnostic, and the total number of questions. Columns marked $D3$ give results similar to $D2$ using the breath just preceding the answer instead of that just following the question. Columns 14 to 16 deal with a comparison of the T and L on the same hour. Knowing that one average was made during L and one during T the assumption is that the L will be algebraically less than the T. If this assumption is correct in a given hour's comparison an "x" is placed in the column, if it is incorrect an "o" is indicated. Column 14 gives these results comparing the $D1$ figures; Column 15 for the $D2$ figures etc. The last column indicates by a cross a correct judgment as to which of the two crimes was L and which T on the basis of blood pressure. The columns are averaged or totalled as indicated and in the case of the totals the number correct reduced to a percentage of the total number.

The results are inconclusive. The final averages of the D's in the two instances show insignificant differences, pointing in the wrong direction if at all. Not over half of the average D's have the proper sign, and the individual D's are correct in only about half of the cases.

It is impossible to judge which method of computing D is the best for none of the three gives more than chance success. There are slight indications that $D3$ is the worst method because column 6 shows a clearer tendency than the others toward a truthful D, and the following column shows that only 31 per cent. of the L questions are correctly diagnosed on the basis of $D3$. This would seem to indicate that after the question is asked and the subject begins to consider his reply, the effect of the lying consciousness dies out rapidly so that during the last breath before replying he is in a practically truthful state of mind.

Neglecting absolute values and attempting merely to select which of the two crimes is L, success is attained in only 2 or 3 of the 7 hours as shown by columns 14 to 16. Nothing depends apparently on the order of the two crimes during the hour for the L came first in 3 of the 7 hours. Of the hours correctly judged the L came first in 2 and the T in the other.

The only suggestion by way of explanation of the failure of series III. is the factor of the habituation, invoked previously with reference to series II. Three of the subjects had already worked in series I. and II. The other subject Q was new and he has the largest negative $D1$ value on L and the largest positive $D1$ value on T in the table. This table, however, is interesting in comparison with that for the next series in which a much different result is obtained when a number of spectators are present. It is to be noted further that the percentage of questions correctly diagnosed is greater in the case of T than in that of L which indicates that the lying consciousness was less frequently present although the person was just as frequently lying.

5. Results of Series IV

The method of Series IV. was identical with that of Series III. except that a number of persons—from 3 to 26—were present watching the subject and trying to determine his guilt or innocence from his reactions. The results are summarized in Table III., which is identical in form with Table II.

Table III

Subject	Lie						Truth						Correct Alternative			Correct B.p.	Correct Jury	
	D_1	No. Cor.	D_2	No. Cor.	D_3	No. Cor.	D_1	No. Cor.	D_2	No. Cor.	D_3	No. Cor.	D_1	D_2	D_3		L	T
A. . . .	−.29	9/13	−.32	7/13	−.24	5/13	.40	5/7	.40	5/7	.33	5/7	x	x	x	x	8/9	2/9
F.——	−.16	5/10	−.06	3/10	−.10	5/10	.04	6/11	.15	6/11	.18	6/11	x	x	x	x		
G.	−.04	5/11	−.01	4/11	.01	4/11	−.02	5/11	−.10	2/11	−.12	2/11	x	o	o	x	3/5	2/5
K. . . .	.05	5/13	−.04	5/13	.05	5/13	.11	9/12	.11	6/12	.07	7/12	x	x	x	x	11/18	7/11
K. . . .	.02	4/9	−.02	2/9	.01	2/9	.13	11/12	.01	5/12	.14	9/12	x	x	x	o	18/26	9/26
K. . . .	−.07	5/8	−.06	6/8	−.07	5/8	.01	5/13	.24	8/13	.04	7/13	x	x	x	x	1/3	1/3
K. . . .	.13	0/7	.17	0/7	.10	0/7	−.04	2/10	.01	3/10	−.12	1/10	o	o	o	x		
M. . . .	−.03	4/12	−.01	5/12	.02	5/12	.04	6/8	.22	6/8	.06	5/8	x	x	x	x	7/11	3/11
N. . . .	.05	2/9	−.03	3/9	.14	3/9	−.02	1/9	.00	3/9	−.01	3/9	o	x	o	x	7/13	4/13
N. . . .	.01	5/12	.03	3/12	.05	3/12	−.14	4/7	−.44	3/7	−.06	4/7	o	o	o	x	4/10	5/10
P.	.03	3/8	−.35	4/8	−.28	4/8	.08	6/8	.23	8/8	.10	4/8	x	x	x	x	1/6	3/6
Av. or total	−.03	47/112 42%	−.06	41/112 37%	−.03	41/112 37%	.05	60/108 55%	.08	55/108 51%	.06	53/108 49%	8/11 73%	8/11 73%	7/11 64%	10/11 91%	60/101 59%	36/94 38%

$D1$: av. of 3 to 5 breaths between question and answer minus av. of 3 to 5 breaths following answer.

$D2$: first breath after question minus first breath after answer.

$D3$: last breath before answer minus first breath after answer.

The two additional columns at the right indicate the number of the jurors who judged correctly on a given crime whether the subject was lying or telling the truth. The first figure indicates the number of persons judging correctly and the second figure the total number of persons. In the two cases left blank there were only one or two jurors present.

The results show a considerable relation between breathing and mendacity. The averages of columns 2, 4, and 6 are all negative while those of 8, 10 and 12 are all positive as required by hypothesis. They likewise vary sufficiently from zero to be significant. If the data of columns 2 and 8 are averaged using the original *D*'s for the individual questions the results are − .04 and .05 respectively and the difference between these two figures is 4.4 times the probable error of difference. Considering however, the number of questions in which *D* is of the proper sign there is no greater correspondence than that due to chance. This agrees with the finding in series I that many question show little effect in the breathing but an occasional one does so to a marked degree and thus alters the average value. Disregarding absolute values and attempting merely to select which crime of the two in an hour is L (*i.e.*, has the average *D* algebraically less) the method is successful in 8 of the 11 hours (73 per cent.) using $D1$ or $D2$, and in 7 hours using $D3$.

A brief investigation was made with this series of the factor of variability. The mean variation was found for each of the individual $D1$ averages that appears in the second and eighth columns of the table. For example for subject *A* the m.v. of his 13 *D*'s whose average was − .29 was found and likewise for his 7 *D*'s whose average was .40. The question was whether these m.v.'s would be consistently greater on L or vice versa with the various subjects. As a matter of fact in this series 6 of them are greater on L and 5 on T while the average of the former is .14 and of the latter .18. Evidently variability of *D* affords no diagnosis. It did not seem worth while to investigate it in other series.

The columns at the right giving the votes of the jury show the unreliability of judging a person's veracity from his directly observable behavior. On the average the judgments of L seem to be correct more often than those of T. This reflects a tendency to be suspicious of the subject in this situation and to give an L judgment too frequently. There is little correspondence between the vote of the jury and the magnitude of *D*. Ranking the jury votes with the one with the largest per cent. voting correctly ranked highest and with the *D*1 values ranked with the largest negative L highest and the largest positive T highest the correlation for L is .58 and for T − .13 using the method of rank differences squared.

The relative merits of the three methods of computing *D* are shown in Table IV.

Table IV

Number of Signs Correct.

	*D*1	*D*2	*D*3
L	5/11	9/11	4/11
T	7/11	8/11	7/11
Total	12/22	17/22	11/22
%	55%	77%	50%
D(T) − *D*(L)	.08	.14	.09

The data is taken from Table III. The first column shows for example that 5 of the 11 *D*1 averages on L's are correct (*i.e.*, negative in sign), while 7 of the 11 *D*1 averages on T's are correct (*i.e.*, positive). The total is 12 out of 22 or 55 per cent. The 0.8 below this represents the difference between the final averages of the *D*1 columns in Table III., *i.e.*, .05 − (− .03). This indicates roughly the total extent to which *D* during L differs from that during T. The other columns give similar figures for the *D*2 and *D*3 methods.

The table shows a tendency for the *D*2 method (*I*/*E* for breath immediately following question minus *I*/*E* for breath immediately following answer) to be the most diagnostic. The value of 77 per cent. as compared with 55 per cent. and 50 per cent., and the corresponding difference of .14 as compared with .08 and .09 point in this direction.

This suggests a return to the data of Series I. to see if a similar relation holds there. It is not worth while presenting the results in detail, but below is reproduced the "average or total" row from Table I., which was based on *D*1. Beneath these figures are given corresponding ones for *D*2.

	Av. *D* for 16 Questions		No. of Questions *D* Correct		No. of Cards *D* Correct	
	L	*T*	*L*	*T*	*L*	*T*
*D*1	−.05	.01	49%	53%	52%	64%
D2	−.11	−.03	56%	47%	65%	38%

From the standpoint of per cent. of questions or cards having the proper sign for *D* there is little difference. In fact as pointed out earlier this aspect is not diagnostic anyway. From the standpoint of average *D*'s there appears a somewhat greater difference between L and T for *D*2 than for *D*1 although *D*2 for T is slightly negative. This agrees then with the finding in Series IV.

This series, then in which the subjects were questioned regarding an imaginary crime in the presence of a jury, but in which no aggressive cross-examination was given, shows a definite correspondence between the breathing ratio and the telling of truth or falsehood. This correspondence is manifest primarily in the average of a group of questions rather than in the number of individual questions correctly diagnosed. It appears that fully as good if not a better diagnosis can be made using *D*2 (first breath after question and first breath after answer) than using the average of several breaths before and after. In the practical application of such methods it will be desirable to compare a record known to be true with a suspected one. In the present case if the problem is to decide between two records one known to be T and the other L, this is done successfully in 73 per cent. of the cases.

6. Blood Pressure

Although the main interest of the experiment was in the breathing, the systolic blood pressure readings that were taken at intervals tend to substantiate previous blood pressure studies[1] and form an interesting basis for comparison with the breathing.

[1] Cf. Marston, W. M., 'Systolic Blood Pressure Symptoms of Deception,' Jour. of Exper. Psych., 2, 1917, 117–163.

In series I. a judgment was made upon each card on the basis of the b.p. reading after the second and fourth questions. These judgments were made without consulting the record as to the actual nature of the recital. No absolutely fixed criterion could be adopted. The two features that most frequently indicated L were (1) rises of 4 mm. or more above what appeared, on the basis of initial and intermediate readings, to be the normal pressure under the excitement of examination and (2) failure of the pressure to descend to normal from the height to which it had ascended during a previous card.

The last two columns of Table I. show the number of cards correctly judged in this first series. The total shows 71 per cent. correct judgments on both L and T. In series II. (same table) where the problem was to decide whether the first four of the eight cards were T or L, no mistakes were made on the basis of the b.p. This problem was naturally somewhat easier for it was possible to observe the pressure over a longer period and there was a greater chance of catching the subject at the height of a lying consciousness. Evidently the b.p. is considerably more diagnostic than the breathing, and habituation in Series II. is not sufficient to overcome the advantage of more extended observation of the pressure.

In series III. and IV. the problem was to tell which of the two crimes on a given hour was T and which L. In Tables II. and III., the columns headed b.p. indicate by x that this judgment was correct and by o that it was incorrect. In Series III. there are 5 correct judgments out of 7—a higher degree of success than in the case of the breathing. In Series IV., 10 of the 11 judgments are correct, 91 per cent.—again somewhat better than the breathing results.

The question naturally arises as to the correspondence between b.p. and breathing. Out of 64 cards in Series I. the two agree in verdict in 35 cases, *i.e.*, 55 per cent. In Series II. and III. there is very little agreement, but it will be recalled that the breathing results are very unreliable in these series. In Series IV. there is agreement in 7 of the 11 cases, — 64 per cent. An attempt was made to correlate the two measures in Series IV. as follows. The minimum b.p. reading on a given hour was subtracted from the maximum. The latter always occurred during the L. These differences were ranked. Similarly the breathing D for the L on a given hour was algebraically subtracted from D for the corresponding T. These values were ranked from the largest positive to the largest negative. These two series of measures were then correlated by the method of rank differences squared. The correlation is $.52 \pm .13$. This indicates a tendency for the person whose b.p. responds readily to the mental state of lying to have his I/E likewise respond readily in the manner suggested by Benussi.

7. Introspection

The subject was asked at the end of each hour to report upon his mental state during the experiment. He was not asked specific or suggestive questions.

Every subject seemed to have some consciousness of deception. This was sometimes described as an emotion of fear (subjects A and C) or as a mere 'guilty feeling localized around the heart' (K). The sensations of quickened heart beat were noticed likewise by B, and there was 'less of this internal disturbance toward the end of the hour.' With N there was also fear not of the total situation but of being inconsistent in replies. N mentioned tension or strain during lies and relief after the lie was over. B mentioned 'excitement like singing in public.' M noted 'some sort of resistance during a lie' and he tried to 'build up a sort of mechanism against it.' All of these factors mentioned point to an emotional state or its bodily accompaniments. It is probably futile to theorize until further work has been done upon the emotions as influencing the I/E, but the following is tentatively suggested. It is possible that fear or excitement on L decreases the I/E while awaiting the signal to reply and that after the response the ratio returns quickly to something like normal. It is also possible that some such factor as concentration of attention on T while thinking about the reply increases the I/E with a subsequent return to normal. This factor would likewise be present on L but overbalanced by the emotional factor. If this were true the I/E after both T and L should be somewhat the same. As a matter of fact such is the case. In series IV. the ratios for the breathing following the replies on T and L differ only about half as much as those preceding the replies. If they are compared further with the normal I/E obtained at the beginning of the hour, the ratios after T are almost identical with the normal and those after L are slightly less than normal, as though the emotional effect had not entirely disappeared.[1]

The subjects often described methods by which they tried to deceive the experimenter and jury if the latter was present. In series I. they were instructed to try to fool the experimenter by their manner. This instruction was unnecessary and was subsequently omitted because the situation developed this set anyway. The usual method especially in Series III. and IV. was to try to keep calm or think of extraneous matters during L and to work up a state of emotion or excitement during T. Subject *M* fixed the lie and then forgot about it as far as possible. Sometimes however, he was unable to banish the consciousness that he was lying, as for example in one instance he began thinking what a good alibi he was telling (12 mm. rise in b.p. at this point). He gave especial concern to his mental state at the times of the b.p. readings. He succeeded several times in raising his pressure by general muscular contraction but this was readily detected. Subject *N* used somewhat similar methods. She tried to think of something else—counted or gazed around the room—during b.p. readings on L and she tried to raise her pressure on T by thinking of personal emotional situations or of someone in whom she had 'considerable interest' (presumably of the opposite sex). *Q* (who had discussed the matter with *M*) was quite successful in controlling his b.p. by 'thinking about his girl' during T, and deceived the experimenter, although it was his first appearance in the ex-

[1] Further work is contemplated upon the I/E during various emotional states and during mental work.

periment. It is to be noted however, that the breathing diagnosis was correct. Subject A tried the same system and raised his b.p. considerably during T, but told an exciting L which raised it still more.

These efforts (sometimes successful) to influence blood pressure by emotional control are quite significant from the practical standpoint. A guilty defendant who knows that the expert is trying to establish a 'normal' record for comparison with the crucial record may perhaps be able to keep that normal record generally high or to induce abrupt rises in it by means of imagining emotional situations. In this way the normal and crucial records may appear similar and the latter be mistaken for a truthful performance. One partial remedy might be to produce greater emotional reaction during the crucial record by more rigorous questioning or by the natural seriousness of the situation, and thus drive the b.p. to a higher point than that to which it could be raised voluntarily. A better procedure would be to use as many other criteria as possible, notably breathing and some aspect of reaction time. While it is possible to imagine an intense emotion during each reading of the blood pressure it is difficult to do this on every question of an examination. Hence the breathing record should be able to catch the normal points in the performance and combined with the b.p. increase the reliability of diagnosis.

8. Summary

The inspiration expiration ratio (I/E) was recorded with an improved technique while the subject was lying (L) or telling the truth (T). The average I/E for the 3 to 5 breaths following the subject's answer was subtracted from the average I/E for the 3 to 5 breaths between the experimenter's question and the subject's reply. The hypothesis under investigation was that this difference ($D1$) should be negative for L and positive for T. In some instances instead of $D1$, the I/E for the single breath immediately following the subject's answer was subtracted from the I/E for the breath immediately following the experimenter's question ($D2$). Systolic blood pressure was measured at intervals.

The material consisted in the first series of cards containing letters and numbers with L and T trials in a random order. The second series was similar except that the L's at a given sitting all occurred in succession and the T's likewise. The third series involved imaginary crimes with the subject fabricating an alibi or recounting one prepared by an assistant. The fourth series was the same as the third except that a 'jury' was present.

Series I. with simple unemotional material shows some corroboration of the hypothesis. The average of *D*1 for all L questions is − .05 while that for all T questions is .04, and the difference between these figures is 4.6 times the probable error of difference. This result is primarily due to a few crucial questions. Many of the *D*'s both L and T are in the vicinity of zero but occasional large negative values are found during L and occasional large positive ones during T. Diagnosis of separate cards in this series (4 questions on a card) is rather unsatisfactory. The results of Series II. are inconclusive, due probably to the habituation of the previous series and the decrease of the deceptive consciousness on similar L cards in immediate succession.

Series III. is likewise inconclusive with regard to the subjects who had performed in the previous series but not with regard to the subject who makes his initial appearance at this point.

Series IV. more nearly approaches practical conditions in material and emotional content. Using *D*1 the grand average difference between T and L is of about the same order as in series I. Using *D*2 it is about half again as large. The effect is here likewise due to a few questions, rather than the majority of questions. In a given 'crime' it is better to note the average *D* than to note what proportion of the questions have a *D* pointing in a given direction. The problem of diagnosing the two 'crimes' covered at a sitting, one known to be T and one L, is solved successfully in 73 per cent. of the cases. This is somewhat the type of problem that would be met in practice where a normal record could be obtained for comparison with a crucial one. The use of *D*2 in this series and to a lesser degree in the first seems somewhat superior to the use of *D*1. Throughout the work there is generally a little greater accuracy in diagnosis of T than in diagnosis of L. This suggests that the subject is not always conscious of lying.

Systolic blood pressure has a greater diagnostic value than the breathing. In the last series it correctly indicates which of the two crimes is L in 91 per cent. of the cases as compared with 73 per cent. for the breathing. The two criteria correspond appreciably more than half the time and if effort is made to quantify the measures the correlation is around .50.

The lying consciousness appears to have an emotional (probably fear) content and it is possible to influence expressive measurements somewhat by emotional control. Some subjects successfully raise their blood pressure during T by imaginary emotions. It is thus important to have other criteria as a check.

The writer feels that the successful application of psychology to the problems of deception will involve a combination of several methods, including the two above described. Suspects examined by such methods would then fall into three groups,—those in which all measures point toward guilt, those in which all measures point toward innocence and those in which the measures disagree. The middle group would contain with present methods less than half the cases, and this number will decrease with gradual improvement of technique.

AN EXPERIMENTAL STUDY OF MOTOR ABILITY

BY F. A. C. PERRIN

University of Texas

In the investigation now reported, an attempt was made to analyze motor ability. The conclusions advanced are based upon the results furnished by a series of complex motor tests, a series of simple motor tests, an intelligence test, the university grades of the subjects, and character estimates of the subjects. The various measurements recorded were compared and correlated. Since the data for each subject cover a wide range, the study was necessarily intensive rather than extensive. Complete records were obtained, however, from a group of 51 subjects, all of them undergraduate students in the University of Texas. The group was selected on the basis of variety in academic and motor accomplishment. The results reported, therefore, may be assumed to be fairly typical.

I. EXPERIMENTAL—METHODS AND RESULTS

A. Tests and Measurements

Each of the subjects was given individually a series of 17 motor tests. A three-hour laboratory period was occupied in giving this series to each individual, and in recording certain other measurements to be described later. The single period, in spite of its length, was thought to be desirable inasmuch as it guarded against such variations in health and in physical tone as would presumably appear during a longer period of time. And it was found that the novelty of the tests provided against boredom and fatigue in a satisfactory manner. Since the same sequence in giving the tests was followed for all subjects, any possible decline in ability due to the duration of the period was made uniform for all. As a matter of fact, most of the subjects, at the end of the period, were inclined to remain in the laboratory room and to experiment further with the tests.

The motor tests consisted of a group of 3, chosen for the purpose of eliciting fairly complex reactions; and of a second group of 14, selected to measure elementary motor functions. A more detailed statement of the distinction between these two groups, and of the meaning of the term, 'motor function,' is reserved for an ensuing section of this paper. The three complex tests were the Bogardus fatigue test, card sorting, and a motor coördination test. The second series of 14 included measurements of reaction time, inhibition of the wink reflex, memory of a motor act, weight discrimination, aiming, the ability to locate points while blindfolded, balancing of various kinds, rhythmic counting, the ability to grasp rhythmic units, tapping, steadiness, tracing, physical strength, and vital capacity. It is obvious, especially in the second list, that the term 'motor' is used in a rather inclusive sense.

1. *Motor Tests*

(*a*) The Complex Motor Tests

1. The Bogardus test was the first of the three given to the subject. Some important alterations in the apparatus were made before it was used. A black wooden cube, somewhat smaller than one of the standard blocks furnished with the apparatus, was glued in the square on which the blocks are ordinarily placed. This served as a platform upon which the subject was forced to place the blocks quickly and accurately as the arms of the machine revolved. The arms, of course, were elevated proportionately. By this device, the task of observing and recording correct reactions was facilitated. While it happened occasionally that a block would fall from the platform, lodge between it and one of the arms, and thus stop the machine, the advantages of the alteration clearly outweighed this disadvantage. In the second place, metal contact points were attached to the revolving vertical tube of the apparatus, and corresponding metal strips fastened to the base, in such a way as to secure automatic registration of the number of revolutions per minute on an electric counter.

These alterations, and the use of an excellent A.C. generator, made possible an almost perfect control over the apparatus.

The subject was given three one-minute tests. His correct reactions were recorded on a hand tally, and their average for the three trials was accepted as the final score. The speed was so adjusted as to demand 54 reactions per minute. A uniform method of explaining the nature of the test, and of demonstrating the technique was followed; but the method of reaching for the cubes and of placing them on the platform was varied, so as to include all legitimate methods of attack. During the actual testing, the task of the experimenter was to observe and record behavior, register successful reactions, and call time. The used of a mimeographed form, the hand tally, and the stop watch made this threefold task not only possible but easy.

The scores of the 51 subjects, in terms of the average number of successful reactions, are given in the following form in order to show conveniently maximum and minimum records, median, modes, and the distribution:

								22										
								22	23	25					30			
	14		16	18		20		22	23	25	26		28	29	30	32		
	14	15	16	18	19	20	21	22	23	25	26		28	29	30	32		
11	14	15	16	18	19	21	21	22	23	25	26	27	28	29	30	32	36	45

The analysis made of the learning behavior of the subjects proved to be extremely suggestive. It had been made evident during the preliminary work that the subjects varied considerably in their methods of meeting the situations presented by the Bogardus test, and that their variations could be classified and described. A rating chart was accordingly prepared for the present experiment. It provided for the scoring of the subject on his ability to place the cubes accurately on the platform, on his skill in picking up the cubes with his right hand and in depositing them with his left hand, on his general quickness with right hand and left hand respectively, on his smoothness of coördination in using right and left hands simultaneously, on the extent to which he looked in the supply box while reaching for the cubes. It also provided for the subject's retrospective testimony, solicited at the completion of the three trials, regarding his awareness of these various items of the reaction, with particular reference to his possible conscious or deliberate attempts to change methods. No intimation of the use of this chart was given to the subject.

He was scored on the first set of items by receiving marks of A, B, C, D, and E, representing decreasing amounts of these different abilities. These marks were standardized: a grade of C under any one heading stood for median performance in that ability, or for median employment of that device. The distribution of these marks was not a matter of difficulty. More E's than A's were recorded, but the C's for each item included the middle group.

The results show that of the 27 subjects who attained the median or who exceeded the median score of the number of successful reactions:

21 were above this median in placing the blocks accurately
19 were above in skill with the right hand
18 were above in coördinating the two hands
16 were above in skill with the left hand
16 were above in moving the right hand quickly
15 were above in moving the left hand quickly
10 were above in looking into the supply box

A correspondence between efficiency in performance and skill in the various aspects of that performance of course was to be expected. These results, however, not only indicate what specific aspects of test behavior count for superiority, but, when compared with the data bearing upon conscious change of method, suggest the possibility of the functioning of an intelligence factor. For these aspects are each theoretically capable of receiving conscious attention and emphasis.

But before the results of these additional tests are presented, reference should be made to changes in method on the part of the subjects during the tests. If their reports are to be accepted, these changes were made, in the main, without antecedent deliberation. And furthermore, in numerous instances, they were adopted unconsciously. By this distinction it is meant, in the latter case, that the individual subject frequently found himself in the act of doing something that he had not been doing before, with no memory of the transition, and with no memory of a preceding plan to effect that transition; while in the former case, it is implied that the subject decided on the impulse of the moment to attempt something new. It should be noted that while these shifts in method of attack generally resulted in increased efficiency, they did not necessarily do so. The fact should be emphasized, too, that a few deliberate and conscious variations were made.

The question of the reliability of such introspective evidence of course arises quite readily. But without entering upon an extended controversy, the writer may state that he is very much disposed to accept this testimony at its face value. The introspection called for was not of the dubious sort dealing with fine and possibly theoretical distinctions among mental processes; in fact, it may be straining the term to call it 'introspection' at all. The subject was merely asked to report about his ideas of concrete acts—a task surely within the powers of an intelligent human adult.

The propriety of labeling these methods of attack as intelligence factors is another matter. They are so designated in this article, somewhat arbitrarily, for reasons given above; the subject can attempt, at least, to vary them. And if these attempts are largely unconscious, they are none the less adaptations which result from experience. Skill in placing the blocks is not (necessarily) an affair of intelligence; but the attempt to employ and control skill is a mode of reaction that can be so regarded.

2. In the card-sorting test, ordinary playing cards were used. The subject was required to sort these by suits into four boxes, each plainly labeled by an extra card. He was instructed to hold the shuffled deck, face down, in his left hand; and at the signal, 'begin,' to pick up the cards, one at a time, with his right hand, look at them and put them into the appropriate boxes. He was told not to correct mistakes.

Three sortings were made by each subject, and the average time, registered with a decimal stop watch, for the three was recorded as his final score. Mistakes were charged up against his final score, each one adding to his total time .02 minute. The method of scoring was explained in advance to the subject. The technique, in general was based upon that used by Brown (1).

The use of playing cards, rather than cards bearing original symbols, was decided upon after due consideration had been given to the fact that previous experience might largely determine the score. It was assumed that the subjects were unequally endowed as far as this factor is concerned—not only for card sorting, but for all other tests. An attempt was made to ascertain the extent to which the subjects indulged in card games, and this inquiry did suggest two rough generalizations: that one who does not play is likely to be poor at card-sorting, and that one who does play is not necessarily good at card sorting. But these are fairly weak tendencies. In the absence of any method of measuring transfer, in any of the tests, it was decided merely to recognize this factor as existing in unknown quantities.

A rating chart, similar to the one used in the Bogardus test, but more simple, was employed. It called for ratings upon the following points: helping with the left hand, moving the deck, throwing the cards successfully (as opposed to placing them), holding the deck near the boxes, using the right hand correctly. A vertical yardstick, fastened by a clamp to the experimental table, enabled the experimenter to measure the distance between the deck and the boxes for each subject; but on the other points the letter rating system was used.

The time scores are given in the following table, each column including a range of .05 minute or less:

						1.09					
						1.09	1.05				
						1.08	1.04				
						1.07	1.03				
						1.07	1.03				.85
	1.34		1.24			1.07	1.01		.95	.90	.84
	1.33	1.28	1.24		1.13	1.07	1.01	.97	.95	.87	.81
1.47	1.32	1.26	1.23	1.20	1.13	1.06	1.01	.96	.94	.87	.81
1.44	1.31	1.26	1.22	1.17	1.11	1.06	1.01	.96	.91	.86	.80

The gradings on the charts establish the fact that the good card sorter must employ certain methods, but they do not prove that these methods alone will guarantee superiority. The efficient individual at card sorting employs his left hand as it is used in dealing cards at a game; throws the cards accurately, and does not attempt to place them in the boxes; picks up the cards with his right hand as he would in dealing; and holds the deck fairly close to the boxes. All of these methods were used by those attaining or surpassing the median score, but they were used also by half of those unable to reach the median rank. Evidently other elements were necessary.

These additional factors were, as one might suspect, speed and the ability to recognize the cards. Instead of attempting to score them on the rating chart, the experimenter tried to measure them by two extra tests. The first, or the speed test, was not used with the present set of subjects, since the test schedule was already over-crowded; but it had been tried out in the preliminary experiment, with a group of 60. It consisted in sorting the cards, irrespective of suit, into the four boxes in order, the remaining conditions of the test being observed. The result was a correlation coefficient of + .80 (P.E., .0605) between this and card-sorting. This coefficient was sufficiently large to establish a tendency for the two activities to correlate.

The ability to recognize the cards (or a corresponding ability) was measured in one of the minor tests—that of reaction time. The Vernier chronoscope was employed, and with it, a set of visual stimuli cards containing single spades, hearts, diamonds, and clubs. Both the simple reactions, and the discrimination and choice reaction tests were given to the subjects. The results of the first yielded a coefficient of + .33 when correlated with card-sorting records; the results of the discrimination and choice test gave + .18. The difference between the two coefficients is probably accidental; together, they would seem to indicate relationship between card sorting and ability to react to card symbols.

The considerations and facts regarding intelligence which were discussed in connection with the Bogardus test logically call for attention at this point. As in the Bogardus experiment, changes in method were observed by the experimenter, and were reported by the subjects as being both conscious and unconscious adaptations. Due to the fact, however, that speed and recognition seemed to outweigh these methods in importance, no attempt was made to evaluate the intelligence factor by a control experiment.

3. The third and last of the complex tests is referred to as the 'coördination test' in this account. It was essentially a test of the ability of the subject to perform disparate activities simultaneously with his two arms. The apparatus consisted of an aluminum plate upon which were secured a slightly elevated triangle and a corresponding quadrilateral, made of hard rubber. The strips forming these figures were of the uniform breadth of 2 centimeters. The outside circumference of the triangle was 54 centimeters, that of the quadrilateral, 48 centimeters.

The subject was given a stylus for his right hand and a similar one for his left. Either stylus, upon coming in contact with the aluminum plate, would register a sound with an electric buzzer. A metronome was used as a time indicator. At the starting signal, the subject began to trace the figures, both hands going counter clock-wise. He was required to start and to stop both hands simultaneously. The tick of the metronome, which was set at 72, was the cue for the right hand to move from one corner of its figure to the next; and for the left hand to do the same. The right hand therefore described its figure (the quadrilateral) in four counts, while the left hand completed its round in three.

In this test, a practice period of one minute was the unit of time. It was followed by a minute of rest. The criterion of successful accomplishment was the ability to trace both figures without error for 10 successive rounds, with the quadrilateral as the counting standard. The number of one minute practice periods required for this was the final score.

On the whole, the task was a difficult one. The majority of the learners agreed that it was the most difficult of all the tests, and two of them (both men) passed it up as impossible. For some, however, the coördination came quickly and easily.

The scores show an extremely wide distribution, but with a number of superior scores at the extreme excellent end. In the table, fractional parts of a minute are omitted; and the two subjects who failed are represented by question marks. Since one of them persevered longer than the other, he was assigned, possibly with poor logic, to the 'next poorest' position. The scores are as follows:

																									5			
																									5	4	3	
																			12					6	5	4	3	
						31			23					18	17	16		13	12	11				6	5	4	3	2
?	?	62	56	45	35	31	30	28	23	22	21	20	19	18	17	16	15	13	12	11	10	9	8	6	5	4	3	2

No attempt was made at a detailed analysis of this learning task. The subjects varied as much in their comments, and in their accounts of the mental processes involved, as they did in their objective records. While they seemed to agree that voluntary direction of attention was the trick involved in the mastery of the situation, they disagreed as to the efficient method of controlling this factor. For the purpose of the recent experiment, the test was regarded primarily as an objective one.

(*b*) Tests of Elementary Motor Functions

It should be stated at this point that the term 'function' is used merely for purposes of convenience. A criticism of the term will be made in an ensuing section of this paper. The 14 tests listed seemed to include those aspects of any complex motor act capable of theoretical isolation and description. A few tests of apparently identical functions were included to serve as a check upon the assumption just made.

1. The reaction time of the subject was measured by the Vernier chronoscope. Two kinds of visual stimuli were used —colored discs, and playing card symbols. Both were employed in each of two types of reaction called for. These symbols were employed in each of two types of discrimination-

and-choice reaction. Four series of stimuli, therefore, were given.

The inter-correlations existing among the four series, and the correlations between these and the results of other tests, will be presented later. The following condensed table discloses the tendencies found in the results:

	Longest Reaction Time	Shortest Reaction Time	Median Reaction Time
Simple reaction: colors..................	460 σ	190 σ	298 σ
Simple reaction: cards..................	474 σ	180 σ	312 σ
Discrimination-and-choice: colors........	598 σ	236 σ	336 σ
Discrimination-and-choice: colors........	540 σ	244 σ	390 σ

2. The wide range of meanings of the term 'inhibition' made the selection of a test of that function, or group of functions, a matter of considerable difficulty. At the same time, inhibitions of some kind presumably appear in any motor learning process. It was thought advisable, therefore, to introduce some test of this nature in the series. The attempt to control the wink reflex seemed to afford an unambiguous example of inhibition in one accepted sense of the term, and that ability was accordingly selected, with some misgivings, as a representative of the function.

The apparatus employed was constructed in the laboratory. A section of a piano movement was altered to suit the purposes of the test, by the addition of a square of thick glass, and by the construction of two cork hammers, which could be made to strike the glass by the operation of two levers. The person to be tested rested his forehead on the glass and fixated upon a designated mark. His position was such that the right hammer would strike immediately in front of his right eye, and the left hammer in front of his left eye.

The subject was given 48 trials. It was explained to him that in any one of these he might be tested by the right hammer, by the left, or by both together. He was scored on one of three points in each trial—on complete success, on partial success, or on absolute failure. These scores were weighted respectively as 2, 1, and 0; and the sum of his weighted scores constituted his final score. As the table of scores indicates, the most successful of the subjects made a score of 85 (out of a possible 96) while eight of them were unable to inhibit at all.

0																					
0																					
0																					
0																					
0		2		4																	
0		2		4												30					
0	1	2	3	4						18	19		21			30	36			41	
0	1	2	3	4	7	10	12	13	17	18	19	20	21	22	24	30	36	38	39	41	44

48					
48	54	58	63	72	85

3. The point has been made that the three major tests as conducted were actually tests of the acquisition of skill. This means necessarily, that in each of the three the factor of habit formation was involved. Since this in turn implies the fixation of movements through repetition, it was thought desirable to measure this function.

A pencil maze with no *culs-de-sac* was constructed for this purpose. It contained 28 segments, with a total length of 313 centimeters. The subject was blindfolded, and was told to trace the path-ways in the maze three times, from entrance to exit, using a blunt rubber tracer. He was instructed, in addition, to memorize the sequence of paths; but he was warned not to do any retracing. The accuracy of his drawing, made from memory at the completion of the third trial, was the measure of his success.

At the best, it can be claimed that this apparatus, as employed, measures the immediate memory for a specific sequence of motor acts. It does not necessarily measure kinaesthetic memory: the subject was at liberty to employ articulatory processes, visual imagery, or any other actual or fictitious mental device. The same can be said, of course, of his methods of establishing habits in the complex tests.

The graphs of the subjects were ranked by four judges, on the basis of the number of segments correctly reproduced, minus the number of segments interpolated in any false fashion. With the exception of the five poorest drawings, which were exceedingly wild and inaccurate, the graphs were easily arrayed in a series of approximately equal gradations of merit. The best graph was an almost accurate drawing of the maze path.

4. The weight descrimination test was given on the rather thin supposition that it measures some phase of motor discrimination likely to enter into more complex processes. In general, the method described in Whipple's 'Manual' was followed. The subject made 12 trials, and gave a judgment for each. The possible range, therefore, was from 0 to 12; but the actual range was from 6 to 12, as the table of scores indicates:

			9		
			9	11	
			9	11	
			9	11	
			9	11	
			9	11	
			9	11	
	7		9	11	12
	7	8	9	11	12
	7	8	9	11	12
	7	8	9	11	12
6	7	8	9	11	12

5. In the aiming test, as in the tracing and steadiness tests, the Stoelting apparatus was used and the general method described by Whipple was followed. Records were obtained for both the right and left hands. The sum, rather than the averages, of the distances between points of intersection on the target blank and pencil points was taken as the unit for scoring, since this permitted the experimenter to rank subjects in order of merit. Since there were 10 points of intersection on the target blank, the figures in the tables below are centimeters and decimal parts of centimeters. Any one score, again divided by 10, would give the average number of millimeters for the individual between pencil point and intersection point.

								4.1	
				8.1	7.0			4.2	
				8.2	7.2	6.0	5.3	4.2	
				8.3	7.5	6.1	5.5	4.5	
			9.0	8.3	7.5	6.2	5.6	4.6	
	11.5	10.4	9.0	8.3	7.6	6.2	5.7	4.8	
12.0	11.6	10.7	9.5	8.4	7.6	6.3	5.8	4.9	3.4
12.3	11.9	10.7	9.7	8.9	7.9	6.3	5.9	4.9	3.5

Distribution of scores—right hand.

20.3 18.8 17.4 16.3 15.5 14.2

				10.0						
				10.0						
				10.2						
				10.2						
				10.3	9.1		7.4	6.0		
			11.3	10.3	9.1		7.5	6.4		
			11.4	10.4	9.3		7.6	6.5	5.1	
			11.5	10.4	9.3	8.2	7.6	6.5	5.3	
			11.5	10.8	9.5	8.5	7.7	6.7	5.5	
13.6	12.2	12.0	11.7	10.9	9.5	8.8	7.7	6.8	5.8	4.6

Distribution of scores—left hand.

6. The aiming test was followed by one in which the subject was directed to register a point with a pencil with his eyes open, and to make a second one, with his eyes closed, as near the first as possible, after having in the interim dropped his hand to his side. The base board used in the aiming test held the paper in place, and the subject stood at the same distance from it observed in the previous test. Three trials were made for the right hand, and three for the left; and the sum of the distances, as before, was recorded as the subject's score.

This test is reminiscent of the ancient parlor game of pinning the tail on the donkey. It was introduced chiefly for the purpose of comparison with the aiming test, since one of the central problems of the experiment was to describe, if possible, the degree of modality of the motor functions.

In the appended table, the scores represent sums of three trials, in terms of centimeters and decimal fractions thereof :

				8.1					3.0
				8.2		6.1			3.1
				8.3		6.1		4.0	3.3
		10.0		8.6		6.2		4.1	3.3
		10.4	9.0	8.7		6.3	5.2	4.2	3.5
12.1	11.1	10.7	9.3	8.8	7.1	6.5	5.5	4.2	3.5
12.2	11.7	10.9	9.3	8.8	7.6	6.6	5.6	4.2	3.8
12.3	11.8	10.9	9.3	8.9	7.9	6.8	5.9	4.9	3.8

Distribution of scores—right hand.

							8.0			5.1		
							8.0			5.1		
					10.1	9.0	8.1		6.0	5.2		
					10.3	9.1	8.2		6.0	5.3		
					10.3	9.3	8.3		6.1	5.5		
					10.7	9.8	8.7		6.2	5.5		3.0
		13.0		11.3	10.7	9.8	8.9	7.6	6.2	5.6	4.7	3.2
17.5	14.4	13.9	12.4	11.7	10.8	9.9	8.9	7.7	6.4	5.7	4.7	3.4

Distribution of scores—left hand.

7. In the balancing tests, five different things were demanded of the subject. He was asked to walk forwards, and backwards, on the edge, ½ inch in width, of a 12-foot board, held up by blocks nailed to its sides; and he was required, in addition, to take a few steps on this board, turn, and walk backwards. He made five attempts to perform each of these tasks. He was then required to balance himself, first on his right foot, and then on his left, on a platform, 10 by 4 inches in size, fastened on the end of a 2″ x 4″ plank, 8 inches long, which was placed on the floor in a vertical position. And finally, he was directed to balance a strip of wood 29 inches in length, 1 inch in width and in thickness, vertically, first with his right forefinger, and then with his left.

For the 'walking and turning' tests, errors were counted; and for the 'balance—standing' test and the 'balancing stick' test, time was counted. The errors were weighted, since the three things to be done were not of equal difficulty. The result of this was a range of from 0 to 90 possible errors. A maximum time limit of three minutes was called for in the balancing tests.

The following results give some indication of the difficulty experienced by the subjects in the various tests:

	Lowest Score, Greatest Number of Errors	Highest Score, Least Number of Errors	Average Number of Errors
Walking and turning tests.......	56	9	32
	Lowest Score, Shortest Time	**Highest Score, Maximum Time (3 Subjects)**	**Average Time**
Balancing—right foot...........	.12 minute	3 minutes (9 subjects)	.45 minute
Balancing—left foot............	.18 minute	3 minutes (4 subjects)	.53 minute
Balancing stick—right hand.....	.03 minute	3 minutes (2 subjects)	.09 minute
Balancing stick—left hand......	.02 minute	3 minutes	.07 minute

8. Inasmuch as the reactions in the Bogardus and co-ordination tests were of necessity rhythmic, it was surmised that tests of this function might possibly be of value. In the first of two tests devised for this purpose, the subject counted aloud with a silent pendulum, while observing it, for five counts. He then, having had opportunity to catch the rhythm, turned, closed his eyes, and attempted to continue counting aloud at the same rate. He was stopped when he lagged behind or forged ahead one count. The number of counts constituted his score.

The scores were fairly well distributed, and were rather widely scattered, as the table shows:

		13			17											
		13	15		17				23	24		26	27		29	30
10	11	13	15	16	17	19	20	22	23	24	25	26	27	28	29	30

32																	
32				37									60				
32	33	34	36	37	40	41	43	46	47	51	53	57	60	66	75	84	93

9. The second of the rhythm tests was designed to measure the subject's ability to grasp rhythmic units. Eight series of dots and dashes were printed upon a sheet of paper, and this paper was fastened on the drum of a kymograph. This instrument was placed in a horizontal position. As the drum revolved, at a slow rate, and carried the dots and dashes constituting one series past a stationary mark, they were translated into sounds made by an electric buzzer. The subject, after hearing one complete series, gave his verdict of

'yes' or 'no,' according to his judgment of whether or not the series of sounds came in rhythmic groups. Four of the series were arranged to show these groups, and the number of units in these four varied from 5 to 9. The range of possible correct judgments therefore was from 0 to 8. While this possibly permitted too much of the chance factor, the subjects on the whole were quite certain of their judgments. The scores were distributed as follows:

			5			
			5			
			5	6		
			5	6	7	
			5	6	7	
			5	6	7	
			5	6	7	
			5	6	7	
			5	6	7	
			5	6	7	
		4	5	6	7	8
		4	5	6	7	8
2	3	4	5	6	7	8
2	3	4	5	6	7	8

10. The Burrough's Calculator was substituted for the tapping board in the speed test. This instrument is at least more foolproof than the board. The subject held a pencil, rubber end down, and used the lower, right-hand '1.' He was given three trials of .10 minute each, and his average for the three was his score. Right-hand and left-hand records were obtained, as follows:

						40											
						40			43	44							
						40			43	44							
						40			43	44			47				
				38	39	40	41		43	44			47	49			
		35	36	38	39	40	41	42	43	44	45	46	47	49		51	
29	32	35	36	38	39	40	41	42	43	44	45	46	47	49	50	51	53

Distribution of scores—right hand.

											42							
								39			42							
								39			42							
								39			42							
								39	40	41	42							
				35			38	39	40	41	42		44					
			34	35	36		38	39	40	41	42	43	44	45	46	48	49	
30	31	33	34	35	36	37	38	39	40	41	42	43	44	45	46	48	49	51

Distribution of scores—left hand.

11. The subjects were tested for steadiness, both for the right hand and for the left. The Stoelting apparatus was used, and the Whipple directions followed. Since time was recorded with the decimal stop watch, the scores read in terms of minutes and decimal parts of a minute:

														.27		
				.13	.14	.16					.23		.26	.27		.31
.06	.08	.09	.11	.13	.14	.16	.17	.19	.20	.21	.23	.24	.26	.27	.30	.31
					.45											
.32	.33				.45								.74			
.32	.33	.34	.39	.44	.45	.46	.48	.50	.55	.60	.69	.70	.74	.80	.98	1.01
1.32	1.77	2.73														

Distribution of scores—right hand.

		.08	.09		.12	.13		.15		.18				.26		.28
.05	.07	.08	.09	.11	.12	.13	.14	.15	.17	.18	.19	.23	.25	.26	.27	.28
			.33													
.30	.31	.32	.33	.35	.37	.41	.42	.43	.44	.50	.55	.58	.64	.66	.77	1.03
1.11	1.55	2.90														

Distribution of scores—left hand.

12. In the tracing test, again the Stoelting apparatus was used, and the Whipple instructions followed. The board was kept in but one position, however, and the movement of the subject's hand was in one direction only—towards his body. Records were taken for the right and for the left hands. They are given in the table in terms of centimeters:

		33				38										
23.5	28	33	34	34.5	37.5	38	41	42.5	43.5	44	44.5	45.5	47	47.5	48	49
	50		51				55		57		58.5		60.5			
49.5	50	50.5	51	53	53.5	54	55	56	57	57.5	58.5	60	60.5	61	61.5	62
	65															
64.5	65	65.5	67.5	68.5	69.5	70	70.5									

Distribution of tracing scores—right hand.

														50.5		
35				40					48			49.5		50.5		
35	35.5	38.5	39.5	40	42.5	44.5	45	46	48	48.5	49	49.5	50	50.5	51	52.5
	53.5		55													
53	53.5	54.5	55		56.5							63			64.5	
53	53.5	54.5	55	56	56.5	57	59.5	60.5	61	61.5	62	63	63.5	64	64.5	65
66	68.5	71														

Distribution of tracing scores—left hand.

13. Physical strength was tested for the right hand by means of the Smedley dynamometer, and for the back and legs by the back and leg dynamometer. In order to get a uniform distribution for the group of boys and girls, weights were assigned to the girls' records: the difference between the median record of the boys and the median record of the girls was added to the score of each girl. The weighted scores are given in the tables:

							42			47		49	50			54							
							42			47	48	49	50		52	54		56					
							42	44	46	47	48	49	50	51	52	54		56		58		61	
24	31	32	34	35	40	49	42	44	46	47	48	49	50	51	52	54	55	56	57	58	59	61	65

Dynamometer scores—right hand.

							87						95	97					
		77					87						95	97					
		77	80			86	87			90		93	95	97	99				
60	67	77	80	82	85	86	87	88	89	90	92	93	95	97	99	100	101	107	108

	112										
110	112	119	120	124	125	137	158	159	162	205	212

Dynamometer scores—back strength.

												142					
			100			120			130			142					
50	94	96	100	106	110	120	122	128	130	138	140	142	143	148	155	156	158

160																		
160																		
160																		
160																		
160			170	174								202						
160	162	164	170	174	175	176	177	180	188	196	198	202	212	220	225	230	253	260

Dynamometer scores—leg strength.

14. The 'vital capacity' test was the last given to the subject. The wet spirometer was employed, and its cubic centimeter readings recorded. Again the records of the girls were weighted, after the manner described in the preceding test. By this method, 15 was added to the record of each girl. The final results are given in the table:

									43				46								
					38				43			45	46								
					38	39			43	44		45	46	47	48						
	34				38	39			43	44		45	46	47	48		49	51	52	53	
30	34	36	37	37.5	38	39	41	42	43	44	45.5	45	46	47	48	48.5	49	51	52	53	54

2. *Mental Tests*

The Army Alpha test was given individually to each of the 51 subjects. Considered as a test of 'intelligence,' the choice probably was an unfortunate one, since it is not sufficiently difficult for a group of university juniors and seniors. And had the object been an intensive study of mentality, several series of tests would have been more to the point. But we wanted primarily a test of the ability to manipulate verbal and mathematical symbols, without any desire necessarily to identify this with intelligence. The Army Alpha test seemed to meet this requirement; and while the distribution of the scores, given below, shows a departure from the normal frequency surface, the skewness may be charged up, in part, to the small number of people in the group. Each column, it will be noticed, includes the scores covering a range of 10 points.

							169		
							169		
							165		
							165		
						159	165		190
						157	165		189
					146	156	164		188
			128	137	145	155	164		186
			127	137	145	154	162		186
			125	136	144	154	162	179	185
98			120	131	142	153	162	178	180
98	105	113	120	131	142	151	161	174	180

3. *University Grades*

The university grades of our subjects were obtained from the registrar's office, and were transmuted in terms of sigma deviations by the use of Thorndike's tables (2). The values are given in the following table:

											2.78	
											2.73	
											2.72	
							2.39				2.72	2.88
							2.38		2.55	2.67	2.72	2.86
						2.27	2.37		2.54	2.63	2.72	2.85
1.14	1.67	1.78	1.82	2.00	2.15	2.23	2.32	2.47	2.52	2.62	2.70	2.81
							3.79					
2.98	3.17				3.55		3.79					
2.97	3.13	3.28			3.54		3.77					
2.96	3.11	3.23	3.33	3.42	3.52	3.64	3.74	3.89	4.14	4.17	4.43	

4. *Estimates of Character*

Although the plan to secure estimates of character for each of the subjects was part of the original program, the specific items enumerated on the chart were formulated entirely as the result of the observations of the behavior of the subjects during the motor testing period. The chart was mimeographed, and copies of it were given to associates of the subjects. Care was taken to see that the judges did not compare notes before recording their ratings. To the extent that it was possible, the same judges were used for the different subjects; but since not more than half of them, at best, knew intimately more than a dozen of the subjects, this could not be effected with any degree of consistency. All ratings were made under the immediate supervision of the writer. The judges were urged to be quite honest, on the promise that their names would not go on record, and that their unsigned charts, even, would not be seen by anyone except the experimenter.

Five ratings on all items listed were secured for each subject, and the median rating was taken as the final one. For purposes of making the idea clear to the subjects, the traits were arranged by opposites. The specific opposed traits used can easily be criticized; indeed, they were criticized by many of the judges. But they did seem to represent the contrasted traits observed during the laboratory testing. The judges were asked to place the subject in the highest 10 per cent. of University undergraduate students, the next highest 10 per cent. etc., down to the lowest 10 per cent. The complete chart is here reproduced:

Please grade ———————————————————— on the following traits, by placing him in the highest 10%, the second 10%, the third 10%, etc., of University students. Try to avoid either over-estimation or under-estimation. These scores will be used only for purposes of correlation, *i.e.*, individual records will not be used. Place a check (V) to indicate in what percentage group the individual stands in your estimation.

Your judgments will be regarded as absolutely confidential.

	Highest 10%	Intermediate								Lowest 10%	
		9	8	7	6	5	4	3	2		
Poise, self-possession											Embarrassment, resulting in incapacity
Vivacity, energy, 'pep'											Entire absence of such traits
Determination, ability to 'see it through'											Indifferent, willing to let things slide
Practical, common sense, general intelligence											Stupidity, absence of common sense
Aptitude for social or group leadership											Easily influenced by any one
Likes to do things well for the sake of the things themselves											Motives extraneous —money, praise, grades
Sense of rivalry, desire to excel											Perfectly willing to be the most inferior
Practical											Visionary
Deliberate											Spontaneous
Conceited											Depreciative
Optimistic											Pessimistic
Frank, sincere											Posing, insincere
Serious											Frivolous
Graceful, skilful											Awkward, clumsy
Consistent in traits											Variable in traits
Skill in tennis, swimming, etc.											

The ratings for the nine traits which seemed to be of the greatest significance were plotted against the rankings in the Bogardus, coördination, and card-sortings tests respectively. The scatter diagram was used for this purpose, and the actual median of the group represented on the diagram was used. This median, in the case of the character ratings, was rather high, indicating a tendency for these ratings to favor our group of 51 subjects on desirable traits; but as the judges frequently pointed out, this group was a selected one. Approximately 50 per cent. of the subjects were placed in the highest 30 per cent. of desirable traits. This is probably a fair distribution, when we consider the fact that the subjects were of sophomore standing or better, although they were compared, in the ratings, with the entire undergraduate body.

B. Correlations and Comparisons

The principal measurements obtained from the various tests and rating charts were compared and analyzed by three methods: (1) coefficients of correlation among the test scores were ascertained; (2) scatter diagrams for the character estimates and tests were plotted, and their results expressed in tabular form; (3) a table was constructed for the purpose of showing the variability of the subjects in five of the more important measurements. The results of the first two methods are submitted as showing general tendencies, or lack of tendencies, only; the results of the third method may help to explain some of these tendencies.

1. The coefficients of correlation obtained, and their (approximate) probable errors, are given in the following table:

	Bogardus		Card Sorting		Coördination	
	Coefficient	P. E.	Coefficient	P. E.	Coefficient	P. E.
Complex Motor Tests						
1. Bogardus					−.10	.0944
2. Card Sorting	−.06	.0949			+.36	.0834
Simple Motor Tests						
1. Reaction time	+.16	.0932	+.21	.0913	−.06	.0949
2. Inhibition	+.02	.0952	−.11	.0943	+.02	.0952
3. Motor memory	+.02	.0952	+.22	.0908	+.17	.0926
4. Weight discrimination	−.07	.0947	−.03	.0951	−.01	.0953
5. Aiming	+.24	.0898	−.09	.0945	+.12	.0943
6. Aiming—blind-folded	+.29	.0873	+.05	.0949	+.04	.0950
7. Balancing	+.38	.0813	−.06	.0948	−.11	.0943
8. Rhythm—pendulum	+.17	.0926	+.20	.0915	+.29	.0873
9. Rhythm—kymograph	−.02	.0952	+.12	.0943	+.12	.0942
10. Tapping	+.22	.0908	+.02	.0952	+.01	.0953
11. Steadiness	+.09	.0945	−.04	.0950	+.01	.0953
12. Tracing	−.11	.0943	−.01	.0953	+.09	.0945
13. Dynamometer	−.22	.0908	+.20	.0915	+.39	.0807
14. Vital capacity	−.14	.0939	+.09	.0945	+.22	.0908
15. Army Alpha test	+.03	.0951	+.02	.0952	+.10	.0944
16. University grades	−.17	.0926	+.01	.0953	+.21	.0913

2. The results of the scatter diagrams were formulated

in terms of the number of subjects falling below the median in one series of measurements who also fell below in a second series; and of the number who ranked above the median in the two series. This seemed to be the best method of comparing the results actually obtained. With a greater number of cases, and with better correlations, finer comparisons could be made. An attempt was made to discover the critical scores, after Thurstone's method, but this was possible only in a few of the diagrams (3). The results are as follows:

(a) Bogardus and Character Estimates

Of those below the median in Bogardus,		Of those above the median in Bogardus,		
16 (31%)	were below the median;	14 (27%)	were above the median in	Poise
17 (33%)	" " " "	15 (29%)	" " " " "	Vivacity
18 (35%)	" " " "	17 (33%)	" " " " "	Determination
12 (23%)	" " " "	13 (25%)	" " " " "	Common Sense
12 (23%)	" " " "	10 (19%)	" " " " "	Interest
11 (21%)	" " " "	13 (25%)	" " " " "	Rivalry
14 (27%)	" " " "	11 (21%)	" " " " "	Deliberation
16 (31%)	" " " "	14 (27%)	" " " " "	Grace
9 (17%)	" " " "	11 (21%)	" " " " "	Skill

(b) Card Sorting and Character Estimates

Of those below the median in Card Sorting,		Of those above the median in Card Sorting,		
17 (33%)	were below median;	17 (33%)	were above median in	Poise
14 (17%)	" " "	12 (23%)	" " " "	Vivacity
20 (39%)	" " "	19 (37%)	" " " "	Determination
8 (15%)	" " "	8 (15%)	" " " "	Common Sense
17 (33%)	" " "	16 (31%)	" " " "	Interest
14 (27%)	" " "	15 (29%)	" " " "	Rivalry
12 (23%)	" " "	13 (25%)	" " " "	Deliberation
14 (27%)	" " "	13 (25%)	" " " "	Grace
10 (19%)	" " "	9 (17%)	" " " "	Skill

(c) Coördination and Character Estimates

Of those below the median in Coördination,		Of those above the median in Coördination,		
13 (25%)	were below median:	14 (27%)	were above median in	Poise
16 (31%)	" " "	14 (27%)	" " " "	Vivacity
10 (19%)	" " "	8 (15%)	" " " "	Determination
9 (17%)	" " "	10 (19%)	" " " "	Common Sense
14 (27%)	" " "	13 (25%)	" " " "	Interest
15 (29%)	" " "	16 (31%)	" " " "	Rivalry
12 (23%)	" " "	13 (25%)	" " " "	Deliberation
12 (23%)	" " "	11 (21%)	" " " "	Grace
16 (31%)	" " "	13 (25%)	" " " "	Skill

(d) Complex Motor Tests

Of those below the median in Bogardus,	9 (17%)	were below median in Cards
“ “ above “ “ “ “	11 (21%)	“ above “ “ “
Of those below the median in Bogardus,	9 (17%)	were below median in Coördination
“ “ above “ “ “ “	10 (19%)	“ above “ “ “
Of those below the median in Cards,	16 (31%)	were below median in Coördination
“ “ above “ “ “ “	15 (29%)	“ above “ “ “

(e) Complex Motor Tests and Army Alpha Test

Of those below the median in Bogardus,	11 (21%)	were below median in Alpha
“ “ above “ “ “ “	12 (23%)	“ above “ “ “
Of those below the median in Cards,	15 (29%)	were below median in Alpha
“ “ above “ “ “ “	14 (27%)	“ above “ “ “
Of those below the median in Coördination,	15 (29%)	were below median in Alpha
“ “ above “ “ “ “	13 (25%)	“ above “ “ “

(f) Complex Motor Tests and University Grades

Of those below the median in Bogardus,	10 (19%)	were below median in grades
“ “ above “ “ “ “	10 (19%)	“ above “ “ “
Of those below the median in Cards,	15 (29%)	were below median in grades
“ “ above “ “ “ “	14 (27%)	“ above “ “ “
Of those below the median in Coördination,	17 (33%)	were below median in grades
“ “ above “ “ “ “	16 (31%)	“ “ “ “ “

3. The coefficients of correlation obtained establish either poor correspondence or absence of correspondence between any two series of measurements. The scatter diagrams indicate that a slight positive relationship exists, inasmuch as an individual who falls below median in any one measurement stands a certain chance of falling below in any other series. Neither of these devices, however, gives us information regarding the number of subjects responsible for the low correlations. It is highly important to know whether they were caused by a few subjects showing extreme fluctuations from test to test, or by all of the subjects showing a consistent and uniform variation.

A number of tabulations were made for the purpose of ascertaining the nature of the variations made. The results of two of these may be presented briefly:

(*a*) In one table, the rank of each subject in each of the 19 tests, exclusive of the rating charts, was determined and stated. According to this table, three subjects ranked above the others. The median rank, for each of the three, was 18:

their rankings were above this in 9 of the tests, and below it in the remaining 9. The poorest subject had a median rank of 32. In half of the 19 tests, he ranked below the thirty-second position. The distribution of these ranks was fairly symmetrical. Evidently, when all of the measurements are taken into consideration, the fluctuations in the record of any one subject are extreme.

(*b*) In this table, however, it is assumed that all tests, mental and motor, are of equal value. But it is obvious that some of the complex tests, although not necessarily more difficult than the simpler ones, clearly are of greater significance so far as the kind of ability which they call for is concerned. Accordingly, a selection was made of the five tests which seemed to measure the most highly organized forms of behavior ability, and the records of all of the subjects were compared, with reference to their positions above and below the median lines. The five selected were the Bogardus, card sorting, and coördination motor tests, the Army Alpha test, and university grades. The results are as follows:

(i) None of the 51 subjects were above median in all of the five.

(ii) 11 of the subjects were above median in 4 of the five measurements. Each of the 11 subjects, therefore, fell below in one of the 5. The low scores were distributed as follows: Bogardus, 5; coördination, 3; cards, 1; Army Alpha, 1; university grades, 1.

(iii) 14 subjects were above median in 3 of the 5 measurements. Each of the 14 fell below in 2, giving 28 low scores, as follows: Bogardus, 7; coördination, 7; cards, 5; university grades, 5; Army Alpha 4.

(iv) 10 subjects were above in two of the tests. The 30 low scores were due to university grades, 7; coördination, 7; Alpha, 7; cards, 6; Bogardus, 3.

(v) 8 subjects were above in one test only. Among these 8, the following scores obtain: University grades, 8; Army Alpha, 8; coördination, 6; Bogardus, 5; cards, 5.

(vi) 5 of the subjects were below the median in each of the 5 tests.

II. DISCUSSION

The primary purpose of this investigation was an analysis of motor ability. It is sufficiently obvious that any conclusions based upon the results obtained must be regarded as tentative and even speculative. A reason for this statement, even more pertinent than the small size of n, is to be found in the fact that an explanation of motor ability involves a satisfactory account of the entire field of human reactions. A number of interpretations or theories, however, were suggested by the results; and our purpose must be satisfied with an examination of these various possible interpretations, with such an evaluation of each as the evidence seems to warrant.

A brief enumeration of these theories will serve to clarify our discussion. Motor ability may be regarded as (*a*) a general factor or function (*b*) the resultant of a number of specific functions, operating simultaneously (*c*) an ability based upon a few general modes of motor reaction (*d*) a function of intelligence (*e*) a response which is determined by more general traits of character, rather than by motor habits or abilities (*f*) a complex reaction which does not permit analysis in terms of functions or general methods of response. While this list of possibilities is fairly comprehensive, it does not extend beyond the interpretations suggested by the data.

A. It seems quite patent that motor ability is not general, but that it is somewhat definitely specialized. As a statement of fact, this means merely that absence of correlation obtained among the scores of the various tests, both complex and simple. Similar results were obtained long ago by Wissler (4). But this investigator was not especially interested in the analysis of motor ability, and his tests consequently were not primarily motor. To the extent that the two investigations are comparable, the results of Wissler are corroborated by the present findings.

To this statement of fact must be added one of explanation. A factor of general motor ability may exist; but if it does, it is lost in combination with other (non-motor) factors as the individual proceeds from one motor test to another. The futility of attempting to predict the degree of success in the coördination test from the record established in the Bogardus test, fully demonstrated in the absence of correlation between the two, was made quite plain to the experimenter before the results were tabulated.

The analogy with the concept of general intelligence is of interest in connection with the foregoing conclusion. Of course, we could pool results, as psychologists have done with mental test results, and assume that the final score of the subject would represent his general motor ability. And we would compare the A.D.'s in motor test scores with the A.D.'s in mental test results. But if this procedure resulted in a marked difference in the two sets of A.D.'s, it would show merely that a factor, or general set of factors, is more operative in the one than in the other; it would not demonstrate the presence of a single general factor, either of intelligence or of motor ability. Undoubtedly, motor tests could be selected so homogeneous in character that high correlations would obtain among them. Varieties of card-sorting tests would be a case in point. Our motor tests however, were selected on a priori grounds, with the end in view of testing each of the possible (assumed) motor functions, while mental tests are frequently assembled on the basis of high correlations, either among themselves, or with school grades. In our conclusion, then, it is not implied that motor ability is to be distinguished from mental ability by the fact of its greater variability.

B. Our data, on the whole, argue against the theory that a complex motor performance can be explained in terms of a number of unit motor functions. It would assuredly simplify psychological doctrine if we could so explain complex motor ability as an agregate, or as a resultant, of elementary motor habits, reflexes, or the like. This would give us an atomic concept which would permit easy analysis of complicated responses, expressed graphically by means of a motor function profile, or some similar device. But motor skill is seemingly not so simple.

This conclusion is based, first of all, upon the fact that no complex test correlates sufficiently with any simple test, or group of simple tests, to justify such an assumption. An attempt was made to apply the method of partial correlations, with the scores which promised reward for the work involved, but no positive results were obtained. The differences ascertained among the individual coefficients of correlation are small—so small that, in the majority of cases, we may believe them to be chance differences. It is true that differences extending over 25 points (about twice the P.E. of the difference between two coefficients) may be assumed to be of significance; but when we note the coefficients which are relatively high, we find that the tests which yielded them fail to show theoretical relationship. Thus, among the relatively high coefficients are the following: motor memory and card sorting; balancing the body and Bogardus; physical strength and coördination; speed and Bogardus. Of these, only the last named seems to suggest logical relationships. And, although we might reasonably expect an actual relationship between card-sorting and speed, coördination and tracing, none was made evident by the coefficients. Even card-sorting, and the discrimination-and-choice reaction test in which card symbols were used, gave a coefficient of only + .18. In fact, the distribution curve of the coefficients was proved to be approximately normal, suggesting that the differences were due to chance factors.

Even more to the point is the question of the existence of simple motor functions. Notwithstanding the fact of objective similarity in the various manifestations of speed, of rhythm, of balancing, and of reaction, it is doubtful if these rubrics represent true functional entities. We can believe, of course, without question, that physical strength is, generally speaking, physical strength; and that a capable piano mover might qualify at the task of loading railroad ties. But it does not follow that the speed required in base ball is the speed demanded in typewriting, or that the rhythm necessary in dancing is that involved in the coördination test.

In connection with this conclusion, the following coefficients are of interest:

Simple color reactions and discrimination-and-choice color reactions, + .59
Simple color reactions and discrimination-and-choice card reactions, + .40
Simple color reactions and simple card reactions, + .33
Locating a point (Test 6) and aiming, + .26
Speed in tapping, and speed in simple color reactions, + .06
Rhythm (units) and Rhythm (counting), + .04

The principle of identical elements is suggested by the above list of coefficients. But this principle should not be confused with assumption that speed, for instance, is a function which can be measured and tested. The criticism may be made that the simple motor functions here described were not the ones which actually constituted the three major test reactions, or that the tests employed were not adequate tests of the functions operative. The ones described, however, seemed at the time to enter into the complex performances. They are, moreover, the ones ordinarily described in our literature. And it is possible that some of the 'simple' functions were, in reality, as complex as any of the three major tests. We can say merely that they were 'simple' in terms of the number and scope and variety of observable movements involved. Only to the extent that these movements can be transferred to a number of other tasks can they be regarded as functions.

C. A third explanation remains, still within the field of motor reactions. It is conceivable, at least, that an analysis of motor skill will disclose, not a general function, and not a sum of specific functions, but a series of general modes of reaction to be characterized as longitudinal rather than transverse. Thus, we may assume a common ability in gross reactions as opposed to fine; or uniform skill in a series of tasks depending upon the amount of ambidexterity involved; or an aptitude in tests involving some form of locomotion as over against some form of hand manipulation. The list, of course, could be extended considerably.

Unfortunately, our experimental data do not contribute directly to an evaluation of any theory of this type. Some attempt was made to isolate the factor of ambidexterity, but the evidence obtained was not conclusive. It does show, however, that the degree of ambidexterity possessed by an individual seems to vary from test to test. This particular explanation principle, therefore, promises little. Again, a test involving fine coördinations (the game of jack straws) was given to a group already tested with the Bogardus apparatus, but the results failed to indicate group differences. The explanation suggested, then, must be regarded as hypothetical only.

D. According to a fourth theory, motor ability may be regarded as a function of intelligence. All of the tests demanded perceptual analysis of a fairly complex order. Most of them provided for some variation in the method of attack, thereby permitting the subject to exercise a degree of choice in his reactions. Moreover, since successive scores were recorded, each test assumed the proportions of a learning process; and each involved, in consequence, a play of selective and retentive factors. Verbal processes, analytical as well as expletive, were elicited in abundance from the subjects. Theoretical considerations undoubtedly justify the supposition that intelligence may play a determining rôle in test reactions. Our immediate question, then, is concerned with its importance and its kind.

If intelligence is that type of ability which is measured by school grades and psychological tests, it does not function to an appreciable extent in motor accomplishment. Our evidence is again based upon coefficients of correlation. University grades, when correlated with scores in card-sorting, gave a coefficient of + .01; with Bogardus scores, − .17; with coördination scores, + .21. This last coefficient was the highest obtained between University grades and any motor test. The mental test scores showed a corresponding series of relationships. The fact should be noted, however, that the coefficients, although small, were positive; no inverse relationships were established.

These results by no means eliminate intelligence as a determining factor. But it is obvious that if we are to offer this factor as an explanation principle, our justification must be found in the definition given to the term. No common factor of any magnitude appears in the total group of measurements recorded. The assumed intelligence factor, then, which functions in the reactions made to each test, seems to be peculiar to that test; it is specialized and particularized, and the term becomes, as a consequence, descriptive and not explanatory.

E. One of the most prominent features of the test behavior of the subjects was the manifestation of certain traits of character and temperament. These were clearly responses made to the test situations, even if they did include more than the specific motor adjustments elicited. It was thought that the particular set of such responses made by any one individual in the laboratory might be characteristic, in general, of that individual. If so, they might conceivably explain the specific reactions made to the tests, since the habitual general response would necessarily influence any specific response. Our task, then, was to ascertain the extent to which these general responses were characteristic of our subjects. The method employed has been described; but we may emphasize again the fact that the traits listed on the chart were the ones actually observed during the test periods.

The tabulations, based upon the scatter diagrams, would seem to indicate a positive relationship between proficiency in the tests and favorable ratings on the scale. But it should be noted that the method of tabulation employed tends to bring out relationships, since it groups the records into quartiles, while the coefficient of correlation, which takes into account all differences in rank, tends to minimize any actual gross correspondences which might actually exist. As regards positions above and below the medians, ratings and test scores do coincide to a noticeable degree. In order to make the results of the scatter diagram analysis comparable with the coefficients recorded, a number of relationships obtained by the latter method were measured by the scatter diagram method. This comparison shows a very slight tendency for character estimates to agree more closely than test measurements with other test measurements.

Evidently, the rating chart, as used in the present investigation, does not furnish the key for an explanation of motor ability. To reverse the statement, a test of motor skill cannot be used as a diagnostic measure of character and temperament. Perhaps the chief reason for this situation lies in the fact that we cannot assume with assurance the consistent manifestation of such traits in any individual. Poise, definitely manifested on one occasion, may break down on another. Our notes on behavior, as a matter of fact, contain instances of shiftings in emotional tone on the part of several subjects. And even more to the point is the question of the actual existence of the traits. It is possible that the poise manifested on a social occasion may be identical with the poise shown in the laboratory only as regards objective expression: the one may result from certain fixed habits, while the other may be the effect of an inhibitory process. The theory back of the rating chart itself may be responsible for its poor showing.

Yet a certain significance must be attached to these ratings. The median judgment on certain traits of an individual, furnished by five associates of that individual, even when recorded on a chart open to obvious criticisms, serves as a better indication of some forms of motor ability than other forms of ability, either motor or mental. We may reasonably expect then, that with a more adequate chart, better correlations would obtain. This chart would contain headings more closely descriptive of the reactions called for by the tests; and it would provide for variations in the amount of the trait under different conditions. It would be, in fact, a systematic description of such aspects of an individual's behavior as permit generalizations.

F. Although the conclusions resulting from the analysis of the data are seemingly negative in character, they may still contribute to a definite view of motor ability. For one thing, they enable us to question explanations of undue simplicity. In particular, they justify a certain amount of reserve as regards the term 'function.' No one would question the legitimate use of that term as applied to any one motor performance; but when it is assumed that generalized modes of reaction exist, some doubt may reasonably be entertained. Phrenological categories are none the less dangerous when they are expressed in terms of objective behavior.

The net result of our investigation is the suggestion that even the simplest act of motor adjustment involves a play of factors not isolated through standard experimental precautions. One important group of such factors is undoubtedly that of transfer effects. It is useless to imagine that, while transfer may operate in card-sorting, it is completely eliminated in the Bogardus test, merely because the subject may acknowledge his familiarity with cards and his ignorance of the Bogardus apparatus. A second group of factors includes those ordinarily listed in connection with the learning process. The simplest motor test used probably elicited a genuine learning process; and this means that accidental success, for one thing, played an important rôle. That this factor operated may be inferred from the fact that successful and unsuccessful reactions were equally independent of the attempts of the subjects to analyze their problems. A third group of factors are those of the emotional and temperamental variety. Our reasons for not dismissing entirely the evidence furnished by the rating charts were given above. In addition, our notes hint strongly of the function of similar influences, not brought out by the charts. Inferiority complexes, shifting states of embarrassment and elation, stimulating and inhibiting ideas of unknown origin and character, were only too evident. The psychology of motor accomplishment awaits the formulation of a psychology of motivation.

References

1. Brown, Warner. *Habit Interference in Sorting Cards.* Univ. of Calif. Pub., Vol. 1, No. 4.
2. Thorndike, E. L. *Mental and Social Measurements.* Second Edition.
3. Thurstone, L. L. Mental Tests for College Entrance. *Jr. of Ed. Psychol.*, Vol. X., No. 3.
4. Wissler, Clark. The Correlations of Mental and Physical Tests. *Psychol. Monog.*, Vol. III., No. 6.

COMPARATIVE SIMPLE REACTIONS TO LIGHT AND SOUND

From the Psychological Laboratory of McLean Hospital[1]

F. L. WELLS, Ph.D., In Charge
C. M. KELLEY, M.D., Associate
GARDNER MURPHY, A.M., Assistant

1. It is well known that under ordinary laboratory conditions the simple reaction time to sound stimuli is shorter than to light stimuli, though reactions to minimal stimuli seem to be equal (Wundt). Richet's (1) calculations give a mean of 149 for sound and 197 for light stimuli. Recently Dockeray and Isaacs report a similar finding, 155 to 197 sigma, based on series of 25 observations each with candidates for flying service. It also seems that reaction times become shorter as the magnitude of the stimulus is increased. In this case any relationship between sound and light stimulus depends on the relative magnitude of the sound and light stimuli. In Froeberg's (2) results, different relationships thus appear according to the magnitudes of the stimuli selected for comparison. In his subject *W*, who is the present writer, this sound-light ratio (per cent. which the light time is of the sound time) ranges from 90 to 158, with different relative magnitudes of sound and light stimulus. Theoretical equality in magnitude might be established by taking a sound and light showing equal number of just noticeable differences above the threshold, though the difficulties seem all but insuperable. For example, the sound-light ratios in Woodrow's (3) subjects *Ht*, *Vs* and *Sz*, for 'weak' intensities of stimuli, are 111, 139 and 128 respectively, their ratios for 'medium' sound and 'bright' light are 136, 140 and 136 respectively.

2. Apparatus for giving standard stimuli was not available for the present experiments. The sound stimulus was the

[1] As of August 1, 1920.

click of opening an ordinary 20-ohm relay. It is clearly audible, but a weaker stimulus than ordinarily used. Circuit was made through the chronoscope at the moment of impact. The subject is 7 feet distant from the source of the stimulus. The incidence of the stimulus is thus about 6 sigma later than the impact, and this quantity is deducted from the readings.

3. The light stimulus is the darkening of a tungsten bulb by interrupting the current through it. The same interruption establishes the chronoscope circuit, approximately 6 sigma after the interruption. How long after this interruption the dimming of the light becomes effective as a stimulus is uncertain (4). The incidence of stimulus is treated as synchronous with the beginning of timing.

4. The material consists of 33 experiments of 216 reactions each. There are 13 subjects. Eleven subjects, one woman, had only 216 reactions each. The two other subjects, *K* and *W*, had 2,376 reactions each. The term 'A-group' is used to denote the single experiments made with the 11 subjects, combined with the first experiments with subjects *K* and *W*. The single experiment of 216 reactions consists of 108 reactions to sound and 108 reactions to light. Throughout the experiments, small and irregular numbers of sound reactions alternate with small and irregular numbers of light reactions. This was so that practise, or other intercurrent changes might not have undue differential effect on either class of reactions. The subject is forewarned when to expect a group of sound or a group of light stimuli. When 54 reactions of each kind have been taken, there is a pause of perhaps 90 seconds, after which the remaining half of the experiment is completed. The 216 reactions occupy about 40 minutes; the method does not lend itself to taking reactions in very rapid succession. The experimenter's time is economized by distributing the reactions as they are read.

5. The reaction times were measured with a Leeds and Northrup Type H D'Arsonval galvanometer arranged as a chronoscope, in a manner derived from suggestions by Klopsteg (5). This method has been very satisfactory. The

simple reaction observations to date are upwards of 25,000. Of these 7,128 are directly concerned with the present results. The remainder are mostly with pathological cases.

6. The reaction movement was that of releasing, with the right hand, an ordinary telegraph key. In the light reactions of the second experiments with *K* and *W*, correction is made for a constant error approximating plus 107 sigma, occasioned through a tampering with the set-up. The effect on the averages is negligible.

7. The score of each experimental unit (a series of 27 reactions) is the median. Aside from its economy of time, its use is almost compulsory for pathological material and thus for data to be compared with it. The present concern is with the relations of sound time and light time. This is expressed, as in par. 1, by a ratio, the per cent. which the light time is of the sound time. General average of sound time scores for the A-group is 204 sigma, a little less than the average for Froeberg's 16 millimeter drop. The corresponding light time is 237 sigma, some 21 sigma more than for Froeberg's weakest light stimulus. Their ratio as above (termed the 'sound-light ratio') is thus 116 as against 132 for Richet's figures, where the stimuli were apparently farther apart. In the stimuli used by Dunlap and G. R. Wells (6) still greater disparity appears.

8. Individual differences may be said to exist where a number of records taken under similar conditions on different individuals clearly differ more than a similar number of records made with a single individual. The present material compares the results of single experiments with 13 individuals, the A-group, with 11 experiments on each of two individuals, *K* and *W*. The sound-light ratio, *i.e.*, the per cent. which the light time is of the sound time, is as follows in the three groups of data:

	A-group	*K*	*W*
Median	115	134	146
Average	116	133	145
M.V.	7.8	4.5	3.2

9. If the mean diurnal variations in *K* and *W* are representative, the relation of sound time to light time is a function of individual difference, less distinct than the tapping test, and of about the same order of distinctness as the æsthetic and color judgments of a previous study (7). Further consideration is invited by the extreme positions of the ratios in *K* and *W*.

10. One factor to think of is instrumental. The experiments with *K* and *W* were made after those with the other subjects. Alterations were meanwhile made in the set-up. The question resolves into whether the light stimulus could so have been delayed beyond the beginning of timing. Except for the accidental condition mentioned in par. 6, the set-up seems guarded against this. Control calculations from ten records on pathological subjects made with the initial set-up, and ten other pathological records made with the *K* and *W* set-up (4,230 observations), show medians the same in each, 112; av. 112, m.v. 6.5, av. 118, m.v. 11.3 respectively, one outlying case in second group. One patient is in both groups. His ratio in the first is 106, in the second, 103.

11. Some explanation might be sought in the reaction history of the subjects. Those of the A-group, excepting *K* and *W* (all but one nurses) were without special practise so far as known. Previous to 1907 *W* had much practise to both sound and light reactions. *K*'s previous practise is negligible, as is also the practise effect in the experiments themselves. The situation seems equally unsuited to interpretation in terms of differential practise. In imagery type, *K* is a visualizer, *W* auditory-motor.

12. If *K* and *W* present idiosyncrasies in this respect, it may be related to the fact that both have always shown rather short reaction times to sound. The physiology of sound and light stimuli, and the complication of the latter with inertia effects, suggest that there may be a relatively lower 'physiological limit' for the sound-reaction. This would account for larger sound-light ratios in those with short reaction times. In the A-group, correlation of sound time with the sound-

light ratio is $\rho = -.52$. Those with quicker time to sound tend towards *relatively* slower time to light.

13. It follows that smaller individual differences should be expected in light than in sound reactions. Between members of the A-group the light variation is about 75 per cent. of that for sound. In Froeberg's work, relative variation of single series also, is notably more for sound than for light. This is the case with *K* and *W* in the present material, but in the A-group there is no difference of this kind.

14. The electrical conditions of the experiment necessitate a get-ready signal. Mechanical control of the prestimulus interval presented problems it was not practicable to solve for these conditions. By listening to a stopwatch, the operator kept the prestimulus interval in half the cases, irregularly interspersed, as close to one second, in half the cases as close to three seconds, as he could. It is possible that this necessity is a virtue (8). Special measurements indicated the following properties for the prestimulus intervals:

Length in seconds of the prestimulus intervals with the two experimenters.

	"One Second"		"Three Seconds"	
	Av.	M.V.	Av.	M.V.
Experimenter *K*	1.21	.09	3.10	.18
" *W*	1.20	.11	3.22	.14

These are the intervals which in further allusions will be known as the 'one-second' and 'three-second' intervals respectively. *W* was the experimenter in all the present observations except those in which he was subject.

These prestimulus intervals make no difference in the averages of the A-group, or with *W*. The sound-light ratios are there practically equal, with either prestimulus interval. In the case of *K* it makes a distinct difference, and so, that the light stimulus is relatively favored by the longer prestimulus interval. His figures are as follows:

Sound-light Ratios in Subject *K*	Av.	M. V.
One second prestimulus, first half of experiments	142	8.6
One second prestimulus, second half of experiments	136	7.1
Three seconds prestimulus, first half of experiments	131	5.8
Three seconds prestimulus, second half of experiments	126	5.7

15. To sum up: The relation between reaction-time to sound and light is dependent on the magnitude of the stimuli compared. An absolute sound-light ratio might be formulated only on the basis of stimuli having the same sensation value for different subjects. Comparison of the reaction times of a relatively fixed sound and light stimulus, shows individual differences in their respective values for the sensation-reaction process. Subjects with relatively short sound reaction times do not have light reaction times decreased in the same degree. Correspondingly, there is somewhat more individual difference in reaction time to sound than to light. Idiosyncrasy in the sound-light ratio is observed, in respect to special favoring of sound, and in respect to the effect of different prestimulus intervals.

REFERENCES

1. Richet. Dictionnaire de Physiologie, 1898, 3, 19.
2. Froeberg, S. The Relation between the Magnitude of the Stimulus and the Time of Reaction. *Arch. of Psychol.*, 1907, No. 8, pp. 38.
3. Woodrow, H. Reactions to the Cessation of Stimuli and their Nervous Mechanism. *Psychol. Rev.*, 1915, 22, 423–452, esp. p. 432.
4. Johnson, H. M. The Influence of the Distribution of Brightnesses over the Visual Field on the Time Required for Discrimination Responses to Visual Stimuli. *Psychobiology*, 1918, 1, 459–494, esp. p. 468.
5. Klopsteg, P. E. A New Chronoscope and Fall Apparatus. Journ. Exp. Psychol., 1917, 2, 253–263.
6. Dunlap, K., and Wells, G. R. Experiments with Reactions to Visual and Auditory Stimuli. *Psychol. Rev.*, 1910, 17, 319–336.
7. Wells, F. L. On the Variability of Individual Judgment. Essays in Honor of William James, 1908, 511–549.
8. Woodworth, R. S. Elements of Physiological Psychology. 1911, p. 481. Quoted by Todd, J. W., Reactions to Multiple Stimuli, *Arch. of Psychol.*, 1912, 25, p. 11.

AN EXPERIMENTAL STUDY OF THE "STIMULUS ERROR"

BY SAMUEL W. FERNBERGER

University of Pennsylvania

Titchener has clearly made the distinction between the stimulus attitude and the process attitude in the formation of sensory judgments.[1] In the case of the process attitude, the observer is instructed to judge process only and, indeed in most cases, only one attributive aspect of a given process. This, Titchener insists, is the only proper psychological attitude for experimental work and is obtained only with observers, trained in introspection, who have been given very carefully worded instructions so that the process to be reported is absolutely unequivocal. Under the stimulus attitude, on the other hand, the observer judges not process, but the stimuli themselves. This, Titchener insists, is therefore not psychological, but is rather on the meaning level and gives judgments of an equivocal sort. Hence he terms it, not the stimulus attitude, but the 'stimulus error.'

The equivocality of the judgments under the stimulus attitude is well shown for the case of the two-point cutaneous limen by Titchener in a later paper.[2] He here points out that in this sort of an experiment there exist several perceptive forms between exact 'one-ness' and clear 'two-ness.' Under the process attitude all of these intermediate forms are reported, while under the stimulus attitude, all of these have the meaning of 'more-than-one-stimulus,' and are therefore

[1] Titchener, E. B., 'Text-Book of Psychology,' 1910, pp. 202f., 218.

[2] Titchener, E. B., 'On Ethnological Tests of Sensation and Perception with Special Reference to Tests of Color Vision and Tactile Discrimination Described in the Reports of the Cambridge Anthropological Expedition to Torres Straits, *Proc. Amer. Philos. Soc.*, 1916, **55**, 204–236. For a discussion of the equating of stimulus attitude with equivocal determination, *cf.*, Boring, E. G., The Control of Attitude in Psychophysical Experiments,' *Psychol. Rev.*, 1920, **27**, 440–452.

reported as 'two.' Hence, under the stimulus attitude, the two-point limen will be smaller and the observer will appear to have a greater sensitivity.

Curiously enough the discussion on these points has been largely epistomological and very little experimental work bears directly on the distinction. Titchener bases his discussion of the Torres Straits results on the experiments of Gates[1] and of others who determined the intermediate forms. These experiments were performed from another viewpoint and in more detailed fashion by de Laski and later by Friedline.[2] These investigators, both working with the two-point limen, take into account the intermediate perceptual forms between 'one-ness' and 'two-ness' and emphasize the importance of such a consideration in working out the statistical limen for duality.

The use of lifted weight stimuli presents a situation in which the distinction between the two attitudes should be important. All of the earlier work in this field has been performed under the stimulus attitude—the observers have been instructed to judge the differences between the weights themselves. In the case of lifted weights, it has been shown that the processes involved in the formation of the judgment are exceedingly complex and that these processes actually vary from individual to individual observer and for the same observer at different stages of practice.[3] It has been found that pressure sensations on the tips of the fingers, kinæsthetic sensations localized in the hand, wrist and forearm might all be involved. In the early stages of practice, the pressure criteria seem to be most used but, after considerable progressive practice, most observers tend to form their judgments on the basis of kinæsthetic criteria. Working under the stimulus attitude, the observers seem to use a cri-

[1] Gates, E. J., 'The Determination of the Limens of Single and Dual Impression by the Method of Constant Stimuli.' *Amer. Jour. of Psychol.*, 1915, **26**, 152–157.

[2] de Laski, E., 'On Perceptive Forms Below the Level of the Two-Point Limen,' *Amer. Jour. of Psychol.*, 1916, **27**, 569–571. Friedline, C. L., 'The Discrimination of Cutaneous Patterns Below the Two-Point Limen,' *Amer. Jour. of Psychol.*, 1918, **29**, 400–419. A description of the earlier experiments will be found in Friedline's paper.

[3] Fernberger, S. W., 'An Introspective Analysis of the Process of Comparing.' *Psychol. Mono.*, 1919, No. 117, pp. 161.

terion of higher order and one which is on the perceptual level—that of the second weight going up 'higher' or 'faster,' 'lower' or 'slower' than the first weight—a perceptual criterion, probably compounded of temporal, intensity and extensity (of localization) attributive kinæsthetic components.

Here then would seem to be an excellent field in which to discover whether the judging under stimulus attitude or the judging of some particular attributive process will have any effect upon the statistical limens. This problem has been recently attacked, from one point of view, by Friedländer.[1] He employed lifted weights as his stimuli and, by means of instructions, set up the following different attitudes in his observers: (1) An attitude of judging 'pressure sensations on the inner side of the fingers,' (2) an attitude of judging strain or 'Kraft' sensations, (3) a general attitude of judging the fused perception representing the stimulus (*cf.* pp. 134f.). The lifting was performed for all three series with and without gloves in an effort to ascertain if this change in the objective conditions had any influence on the frequency of the different judgments for the different series. Friedländer finds that, in the pressure attitude series, the presence of a glove on the lifting hand has a great influence on the judgments, while it has little influence in the other two series. He also compares the sensitivity under the objective and subjective attitude and concludes (p. 191) that the sensitivity under the objective attitude is somewhat greater than under the subjective.

Friedländer has shown, then, that it is possible to have the observer adopt either an objective attitude or an attitude of judging only one of several elements in a complex, and, that the adoption of one or other of these attitudes may effect the threshold values. It would appear, however, that his instructions for the subjective attitudes (especially the kinæsthetic) were still too indefinite and equivocal and were therefore capable of further refinement. Also his statistical conclusion regarding the effects of the attitudes upon the

[1] Friedländer, H., 'Die Wahrnehmung der Schwere,' *Zeit. f. Psychol.*, 1920, 83, 129–210.

measure of sensitivity are not conclusive, inasmuch as he reports in this connection the results of only one observer whose reactions were made upon comparison stimuli with too wide an interval (we believe) and who gave too few judgments for them to be of great statistical value.

It seemed then worth while to repeat these experiments with more exact and less equivocal instructions; to use more observers and to use observers at different stages of practice; and to obtain more judgments on a more appropiately arranged series of stimuli.

The stimuli consisted of a standard weight of 100 grams and comparison weights of 88, 92, 96, 100 and 104 grams. Hard rubber stimuli were used so as to eliminate the temperature sensations.[1] These weights were presented in the first time order: *i.e.*, with the standard stimulus always presented first, and with the space error eliminated. A double series of weights were presented in a carefully arranged order so as to eliminate, as far as possible, the effects of the order of the series.[2] The manner of lifting has been elsewhere described in detail.[3] The only difference between the manner of lifting in the earlier experiment and in the present one is that, in the former, the lifting was continuous and the time between the pairs was kept constant while, in the present experiment, the observer was allowed to stop between each pair if he so desired.

Three observers took part in the experiment. Mr. Carroll C. Pratt, A.M., assistant in psychology (Ob. *P*) had a great deal of practice in experimentation of this sort and under these objective conditions, having taken part in a similar experiment last year. Miss Marjory Bates, A.M. (Ob. *B*), graduate student in psychology, had a good deal of introspective training and had some little practice in lifted weight experimentation in the laboratory course last year. Mr.

[1] Fernberger, S. W., 'A New Form of Stimuli for Lifted Weight Experiments,' *Amer. Jour. of Psychol.*, 1920, **31**, 147–151.

[2] Fernberger, S. W., 'Interdependence of Judgments Within the Series for the Method of Constant Stimuli,' Jour. of Exper. Psychol., 1920, **3**, 126–150.

[3] Fernberger, S. W., 'On the Relation of the Methods of Just Perceptible Differences and Constant Stimuli,' *Psychol. Mono.*, 1913, No. 61, 6–11.

Henry M. Halverson, A.M. (Ob. *H*), graduate student in psychology, was totally unpracticed in this sort of experimentation but had a certain amount of introspective training. The experiments were performed in the laboratory of experimental psychology, Clark University, during the autumn of 1920. All of the observers were given sufficient preliminary practice so that the hand movements became automatic.

Three series of experiments were obtained from each of the three subjects. In the first series (Series *A*) the following instructions were given to the observers.

Certain stimuli will be presented to you which will be lifted—in so far as the rhythm and the manner of lifting is concerned—in accordance with your former instructions.

You will report only on the intensity of the pressure sensations on the tips of the fingers. Attend to these pressure sensations only. Three categories of judgment are permitted—'greater," less' and 'equal.' A greater or a less judgment indicates that the intensity of the pressure sensations on the tips of the fingers aroused by the second stimulus, was greater or less respectively than the intensity of the pressure sensations aroused by the first stimulus of the same pair. An equal judgment indicates that the intensities of the pressure sensations aroused by the two members of a pair were subjectively equal.

If you fail to live up to these instructions, either with regard to the categories of judgment or with regard to the formation of the judgment on the basis of the intensity of the pressure sensations only, you will not give a judgment but will indicate the fact by reporting the word 'failure.'

Intensity of pressure sensations in the above is to be taken in the sense of mental process only.

The second series of experiments (Series *B*) was made under instructions which imposed the judging of the intensity of the kinæsthetic sensations localized in the wrist. These instructions were exactly similar to those for Series *A*, except for the insertion of the words 'intensity of kinæsthetic sensations localized in the wrist' for the words 'intensity of pressure sensations on the tips of the fingers.'

The third series of experiments (Series *C*) was made under instructions which were planned to arouse the stimulus attitude in the observers and which are here given in full.

Certain stimuli will be presented to you which will be lifted—in so far as the rhythm and the manner of lifting is concerned—in accordance with your former instructions.

You are to take the stimulus attitude and are to judge the weights themselves. Three categories of judgment are permitted—'greater,' 'less' and 'equal.' A greater or a less judgment indicates that the second weight was greater or less respectively than the first weight of the same pair. An equal judgment indicates that the two stimuli of a pair were of equal weight.

If you fail to live up to these instructions for any reason you will not give a judgment but will indicate the fact by reporting the word 'failure.'

By means of these different instructions it was hoped that the observers would make an attentional abstraction from the mass of processes involved and, in Series *A* and *B*, judge only the intensity of the pressure sensations on the finger

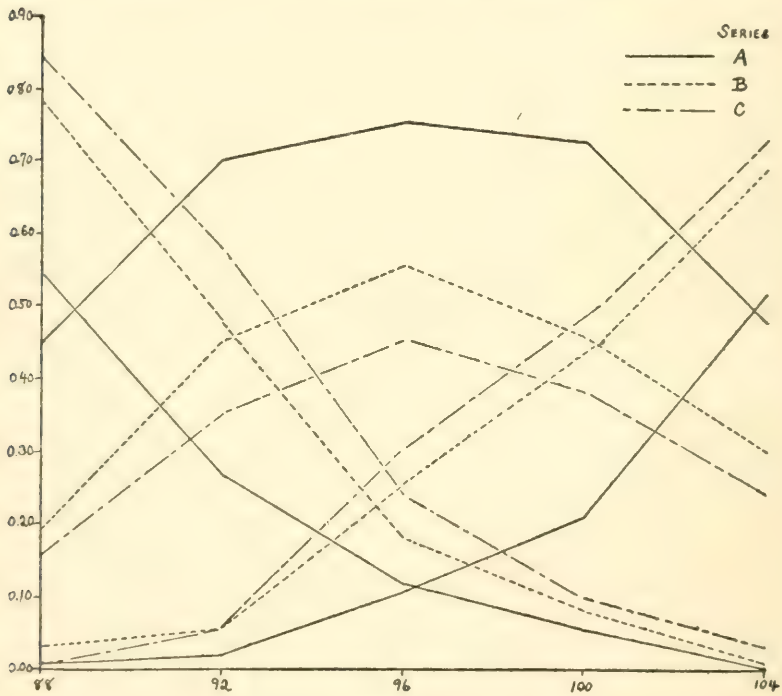


Chart I.
Observer P.

tips and the intensity of the kinæsthetic sensations in the wrist respectively. The statistical values for these series could be then compared with one another and with those of Series *C*, taken under stimulus attitude instructions. Introspective characterizations were taken for every series so that it could be assured that the observers were making the proper attentional abstractions and were forming the judgments on the basis of the processes desired. An examination of these introspective characterizations shows that the instrucions were properly fulfiled.

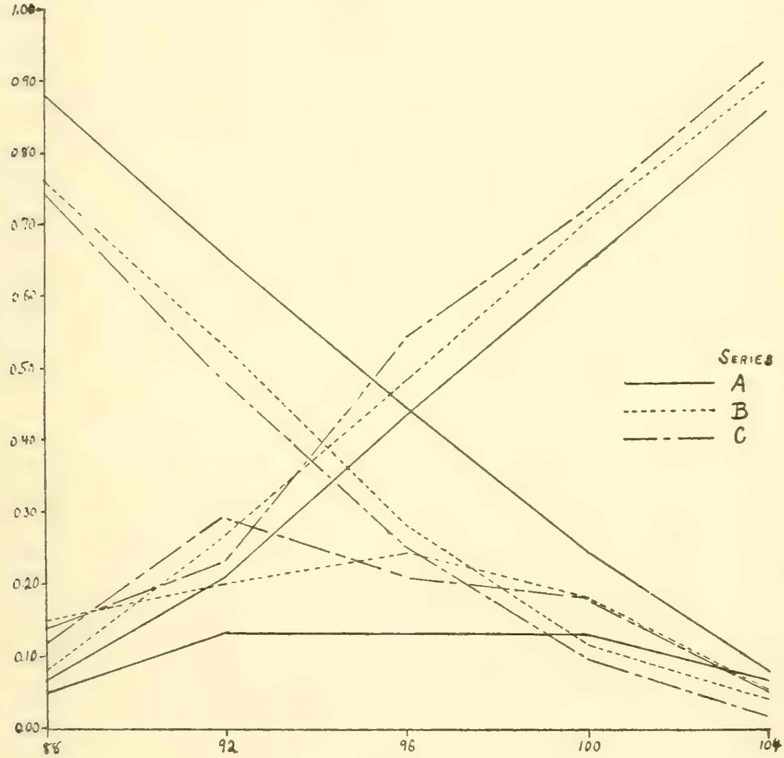


CHART II.
Observer B.

In each series 250 judgments were taken from each observer on each of the five stimulus pairs. The calculations were made in accordance with the Φ (γ) hypothesis by the methods developed by Urban.[1]

The observed relative frequencies for the different series are plotted in the form of curves in Charts I.–III. Each chart contains the curves for all three series for one subject.

[1] Urban, F. M., Hilfstabellen für die Konstanzmethode. *Arch. f. d. ges. Psychol.*, 1914, 21, 337–340.

The curves for Observer P are given in Chart I.; those for Observer B in Chart II.; and those for Observer H in Chart III. The charts are similarly constructed. The values of the comparison stimuli are laid off along the abscissa and the observed relative frequencies are erected as ordinates. The curves for Series A (intensity of pressure sensations on the finger tips) are represented by the solid lines. The curves for Series B. (intensity of kinæsthetic sensations in the wrist) are represented by the short dash lines. The curves for Series C (stimulus attitude) are represented by the alternating short and long dash lines.

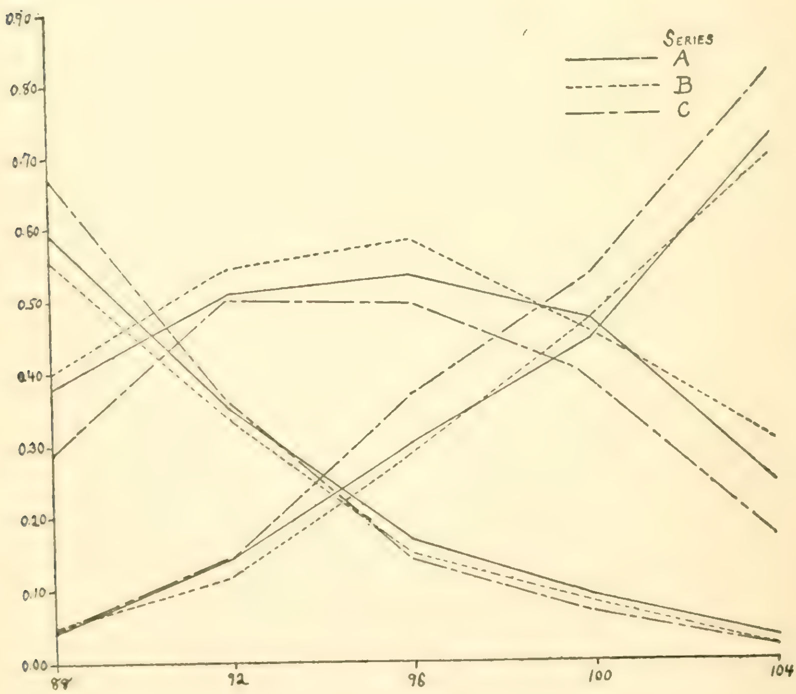


Chart III.

Observer H.

An examination of these charts indicates that all of the curves for every series are of a form which, from inspection, is relatively similar to the form of the Φ (γ) hypothesis. In no case is there a single inversion. This is evidence that each of the instructions given the observers are proper instructions under which to acquire data which may be properly treated, statistically, by the method of constant stimuli. The introspections from the observers gave additional evidence that they were able to make the attentional abstraction required by the instructions and that they were able to form judgments on the basis of them.

It will be observed, however, that these curves vary in form and size for the different series for the same observer—in some cases to a very considerable extent (cf. the equality curves for Ob. *P*, Chart I.). These differences can be best discussed by considering the magnitudes of the constants of the psychometric functions, of the thresholds, the intervals of uncertainty and the points of subjective equality for the different series for each observer.

Table I

Observer *P*

Series	h_1	h_2	S_1	S_2	Interval of Uncertainty	Point of Subjective Equality
A..........	0.102	0.103	88.24	104.49	16.25	96.06
B..........	0.131	0.108	91.83	100.68	8.85	95.84
C..........	0.131	0.117	93.01	99.85	6.85	96.26

Table II

Observer *B*

Series	h_1	h_2	S_1	S_2	Interval of Uncertainty	Point of Subjective Equality
A..........	0.111	0.108	95.08	97.52	2.45	96.37
B..........	0.112	0.109	92.57	96.33	3.76	94.43
C..........	0.112	0.111	91.89	95.80	3.91	93.83

Table III

Observer *H*

Series	h_1	h_2	S_1	S_2	Interval of Uncertainty	Point of Subjective Equality
A..........	0.095	0.097	89.29	100.27	10.98	94.85
B..........	0.093	0.097	88.83	100.49	11.66	94.75
C..........	0.118	0.112	90.26	98.76	8.50	94.35

These values are given in Tables I.–III. Each table is similarly constructed and contains the results of the three series for one subject. In the first column are given the letters of the series. The second and third columns contain the values of the coefficients of precision of the less judgments (h_1) and of the greater judgments (h_2) respectively. The fourth and fifth columns contain the values of the lower (S_1) and the upper (S_2) thresholds respectively. The sixth column contains the values for the interval of uncertainty (S_2–S_1). The last column contains the values for the point of subjective equality

$$\left(\frac{c_1 + c_2}{h_1 + h_2}\right),$$

and indicates the point on the abscissa corresponding to the point of the intersection of the curves of the psychometric functions of the less and greater judgments. In the rows are found the average values for all of the columns for Series *A*, *B* and *C* respectively.

The results for Observer *P* (Table I.) will be first considered. The values for the coefficients of precision for the greater judgments (h_2) are very similar for all three series. The coefficients of precision for the less judgments (h_1) for Series *B* and *C* are identical; while both are considerably larger than that for Series *A*. The thresholds in the direction of decrease (S_1) vary considerably—the lowest being for Series *A*, the highest for Series *C*, and that for Series *B* being intermediate. The threshold in the direction of increase (S_2) for Series *A* is considerably larger than those for either of the other two series; while that for Series *B* is slightly larger than the value for Series *C*. The values of the interval of uncertainty for the different series in *P*'s results are quite different. The interval for the pressure series (*A*) is more than twice as large as that for Series *C* and nearly twice as large as that for Series *B*. Also the value for Series *B* is nearly one third as large again as that for Series *C*—the actual difference being 2 grams. The values, from Observer *P*'s results, for the points of subjective equality are very similar in all three cases—the greatest difference (*B* and *C*) being 0.42 grams.

In the case of the results of Observer *B* (Table II.), the following may be noted. The coefficients of precision for both the psychometric functions of the less (h_1) and greater (h_2) judgments are, for all three series practically identical. The value of the lower threshold (S_1) for Series *A* is considerably higher than those for Series *B* and *C*; while the *B* and *C* values are practically identical. The values for the upper thresholds (S_2) are very similar but that for Series *A* is higher than those for Series *B* and *C*; while that for Series *B* is slightly higher than the value for Series *C*. The values of the interval of uncertainty for Series *A* is smaller than those for either *B* or *C*. The values for Series *B* and *C* are very similar. Similar relations hold with regard to the points of subjective equality for this observer. The value for Series *A* is considerably higher than those for the other series. The differences between the points of subjective equality of Series *B* and *C* is relatively small (0.60).

For Observer *H* (Table III.) the coefficients of precision for the psychometric functions of the less (h_1) and the greater (h_2) judgments are very similar for all three series. The values of the lower thresholds (S_1) are relatively similar for all three instructions—that for Series *B* being the lowest and being very similar to that for Series *A*. The value for Series *C* is somewhat higher than either of the others. The value of the upper threshold (S_2) for Series *C* is considerably lower than those for either Series *A* or *B*, while the values for *A* and *B* are almost identical. The value of the interval of uncertainty for Series *C* is considerably smaller than those for the other two series. The values of the interval of uncertainty for Series *B* and *C* are not very different. The points of subjective equality for all three series are quite similar, the greatest difference being only 0.50 grams (*A* and *C*).

This then appears to be a curious situation when one considers the effects of the various instructions for the different observers for the values of the intervals of uncertainty and for the points of subjective equality. These differences cannot be understood, we believe, without also considering the stage of progressive practice in which we find each of the subjects. It has been shown elsewhere[1] that the processes underlying the formation of judgments of small differences in weight taken under the stimulus attitude vary for different stages of practice. For the unpracticed subject, the judgment is made largely on the basis of pressure sensations on the tips of the fingers, while as practice continues, these are gradually lost until there is a complete shift to kinæsthetic criteria.

In the light of these findings only would it seem possible to interpret the present results. Observer *H* was the totally unpracticed observer. For him the series taken under the stimulus attitude gives the smallest interval of uncertainty. The interval of uncertainty for the kinæsthetic instructions is greater than that for pressure. This result might have been expected at that stage of practice inasmuch as pressure seems to be the 'natural' early criterion. Observer *P* was the highly practiced subject and for him the relations are reversed. Again the stimulus attitude series gives the smallest interval of uncertainty. But the introspections of this observer show that, under the stimulus attitude (Series *C*), he judged largely in terms of kinæsthesis and the interval for his kinæsthetic instructions (Series *B*) much more nearly approximates the value for the stimulus attitude (Series *C*) than they do the pressure instructions (Series *A*). The values for Observer *B* are somewhat equivocal. She was also relatively unpracticed and, as in the case of the other unpracticed observer *H*, the interval of uncertainty for the pressure series (*A*) is smaller than for the kinæsthetic series (*B*). Thus far all is as one might expect it to be. But then the value under the stimulus attitude (*C*) is greater than those for either of the other series. This subject states, in her introspection, that the processes for Series *C* were complex—involving tactual and widespread kinæsthetic experiences and a great deal of visual imagery. She also states that the judgments in this series were given with a much lower degree of subjective assurance, on the whole, than in either of the other series. It is possible that this observer was just in the transitional

[1] Fernberger, S. W., 'An Introspective Analysis of the Process of Comparing,' *Psychol. Mono.*, 1919, No. 117, 136–144.

stage between early practice and late practice. Certainly her results show characteristics of early practice inasmuch as the interval of uncertainty for the kinæsthetic series is larger than that for the pressure series. If she were in the transitional stage, it might well be that in the stimulus attitude series, there was a conflict between the partially disregarded pressure criteria, of the early practice, with the newly acquired but not yet perfected kinæsthetic criteria. This would lead to a certain amount of equivocality in the judgments which would show in a greater number of equality judgments—and hence the larger interval of uncertainty—and also in the lack of subjective assurance reported in the introspections.

The values for the points of subjective equality are very constant for Observer *P*, who was the practiced observer. The values for Observers *B* and *H* show a regular tendency to decrease in the order in which the series were taken. This decrease in the size of the point of subjective equality is one of the normal phenomena of progressive practice as has been pointed out elsewhere.[1] The effects are very slight for Observer *H* and very considerable for Observer *B*. Hence it would seem that the results of these two observers, with respect to the value of the points of subjective equality cannot throw light on the present problem, inasmuch as the effects of progressive practice might have cancelled any effects of the instructions. From the results of Observer *P*, it would appear as if the instructions employed had little if any effect upon the value of the point of subjective equality.

Conclusions

1. It is evident that a trained observer is able, by an attentional abstraction and isolation, to judge intensity differences of a single modality from a complex in which several modalities of sensation are present simultaneously.

2. In the case of lifted weight stimuli, the following instructions are capable of fulfilment; namely, judging the intensity of the pressure sensations on the tips of the fingers;

[1] Fernberger, S. W., 'The Effects of Practice in Its Initial Stages in Lifted Weight Experiments and Its Bearing upon Anthropometric Measurements,' *Amer. Jour. of Psychol.*, 1916, 27, 261–272.

the intensity of the kinæsthetic sensations localized in the wrist; and judging, by assuming a stimulus attitude, the relative intensity of the weights themselves. This latter is distinctly a complex perceptual process on the 'meaning' level.

3. All three of these instructions give curves of the psychometric functions which appear to be somewhat similar in form to the Φ (γ) hypothesis.

4. The pressure and kinæsthetic instructions give rise to strictly psychological results. Whether or not the stimulus attitude instructions are strictly psychological or belong to an experimental logic depends on one's definition of psychology. Certainly all of the classical experiments in the field of lifted weights have been performed under this attitude.

5. The coefficients of precision for the psychometric functions of both the less and greater judgments are very similar for all three instructions.

6. The size of the intervals of uncertainty vary for the different instructions. For the unpracticed subjects, the interval for the pressure sensations is smaller than for the kinæsthetic sensations. For the practiced subject, the interval for the kinæsthetic sensations is smaller than that for the pressure sensations. There seems to be a tendency for the interval obtained under the stimulus attitude to be smaller than either those for the pressure or kinæsthetic sensations alone. The stimulus attitude, however, may lead to an equivocal sort of judgment which results in a low degree of subjective assurance; which in turn results in a tendency to increase the number of equality judgments and thus leads to a larger interval of uncertainty.

7. The point of subjective equality does not seem to be effected by the different instructions. The time error is evident for pressure, kinæsthesis and the stimulus attitude. Progressive practice, evidenced by its effect in lowering the point of subjective equality, is evident throughout all the series for the unpracticed observers.

8. This experiment seems to show the futility of properly interpreting the statistical results of a purely behavioristic study without the control of introspective report.

Journal of
Experimental Psychology

Vol. IV, No. 2 April, 1921

THE PURSUITMETER

An Apparatus for Measuring the Adequacy of Neuromuscular Coördination Described together with Illustrative Results[1]

BY WALTER R. MILES

Nutrition Laboratory of the Carnegie Institution of Washington, Boston, Mass.

Introductory

In a recent paper (1) was described a very simple apparatus called 'A Pursuit Pendulum' which is well adapted to measuring the success of an individual's reaction and eye-hand coördination to a moving object. In this case the test is discontinuous. The reactor makes a number of separate trials after each of which there is an interval when he may relax and then readjust himself before giving the next signal for the release of the pendulum. This corresponds to the usual reaction test, measuring as well the coördination following the reaction. To provide a continuous task of this nature, requiring reaction and eye-hand coördination (the natural accompaniment of reaction) has appealed to the writer as very desirable. It has been worked out in the form of an electrical equipment which, although rather elaborate and expensive, produces results which have fully repaid the expenditures. Before entering on a description of the Pursuitmeter and as a basis of judging its merits, it may be well to describe with some detail the ways in which other workers have met the general problem.

[1] In shorter form this paper was prepared for the American Psychological Association, December 30, 1920, and a communication on the subject was made to the Congress of Physiology, Paris, July, 1920.

The usefulness of a continuous test requiring discrimination and choice reaction has appealed to many, and several combinations of apparatus have been devised to this end. As an example, Seashore (2) described in 1902 what he termed a 'Psychergograph' which was developed as a means "(*a*) to call forth a relatively simple and definite complex of mental activity (*b*) to repeat the same for any desired length of time without interruption and (*c*) to measure the amount of work done, the time taken, the quality of the work, and fluctuations in speed and quality." His apparatus consisted of a stimulator and a recorder, the former being a clockwork mechanism carrying a disc on the periphery of which four symbols were repeated over and over in chance order. The symbols were viewed one at a time through a small window. Each reaction operated, by means of electro-magnets, an escapement which exposed a new signal for reaction. The reactor thus made his own tempo. Each of the four reaction keys closed an electric circuit and a graphic record was made on paper tape by four ink pens while a fifth pen provided a time line. When a test had been started all was quite automatic so far as the experimenter was concerned. The pressing of a key, whether it was the right or wrong reaction, changed the symbol, hence aside from reading the reaction time the graphic record had later to be interpreted, when determining the number of errors made, by comparison with the known order of the symbols on the disc. Having used the 'Psychergograph' at Iowa, its many excellent features are personally known to me.

McDougall (3) in 1905 published a description of an apparatus for the study of concurrent mental operations and mental fatigue which has later been modified and is now referred to as 'The Rivers-McDougall Fatigue Machine.' Circular targets, 2 mm. in diameter, are printed in as irregular a manner as possible along the length of telegraph tape. They are spaced at a constant distance of 5 mm. along the length of the paper and given a maximal lateral deviation of 12 mm. A continuous band of such paper is caused, by means of a weight-driven clockwork, to be exposed through a slit. The reactor, with a needle-pointed stylus, tries to spear each target as it is presented to view. The difficulty is increased by quickening the speed, until maximal effort is required from a particular subject. The distance from the prick mark to the target is measured to determine the efficiency of performance. This apparatus has been used by McDougall and Smith (4), Burt (5), M. Smith (6), and others with very good results. The Rivers-McDougall machine is very compact and has the practical advantages of easy portability and independence of laboratory conditions.

More recently McComas (7), among others, has arranged a set-up for recording continuous discrimination reactions which in practice is quite like Seashore's arrangement except it is more quiet in its operation and employs for stimuli four colored lights exposed behind one small ground glass window. A reaction, by pressing the appropriate one of four keys, extinguishes the light which has appeared as a stimulus and causes a secondary electric clock placed opposite a circle of radiating copper contacts to move forward 1/60 revolution and so close a circuit and present the next stimulus. The four stimuli are presented in chance order and there are 60 in the series until it begins to repeat. The wiring arrangement is very ingenious. On the kymograph record which is well condensed one signal records all reactions while a second marker records only correct reactions. When later measuring the time values the errors are thus directly revealed.

A usual part of such apparatus and a seemingly essential feature, as there has not appeared any good way for taking care of the matter otherwise, is the graphic record. Individual reactions, 'single acts,' may be recorded directly from reading the chronoscope at the time of the experiment. Continuous reactions have on the contrary to be measured out in reference to time lines on paper tape or kymograph tracings. There are, doubtless, some instances when it is desirable to measure these reactions individually 'at leisure' and to definitely separate the error score and the time score for the proper analysis of speed, of quality and variability of performance. Practically the final comparison for the individual's working efficiency under this or that condition is usually made on the basis of an 'arbitrary' combination score. It is a question, therefore, if much of critical value would commonly be sacrificed by having the apparatus so arranged that it will automatically combine the error and time scores of all reactions into a total score for the test period. The possible gains of not having to measure the individual reactions, weight the errors, compute averages and otherwise manipulate the data after the experiment, may offset the loss of opportunity for minute analysis.

There are circumstances when it would be valuable, if not indeed essential, to know the results immediately following the test or even to watch them accumulate and sample them by short periods during the test. Dunlap (8), in outlining the practical requirements to be met in testing the resistance of pilots to increasing oxygen deficiency, found that it was necessary to have a single, comparatively brief test, continuous in nature but not in itself fatiguing, and which revealed the results immediately. There were special limitations by reason that the subject must at the same time be connected to a rebreathing apparatus but he concluded that the "final composite reason for using a clinical method (watching the reactions and coördinations of the subject rather than measuring them) comes from the need for rapid work. Graphic methods might be employed but would largely hinder the expedition of the work on account of the time and labor needed for their interpretation." Dunlap assembled on a rigid and especially convenient table a group of stimulus lamps which had to be reacted to when each one lighted by touching a stylus to a corresponding target. When the center of the target was touched a check lamp flashed indicating that the reaction had been correct and within a certain time interval. If the periphery of the target was touched an error lamp flashed. The reactor was to give his first and prompt attention to responding to the stimulus lamps. Two other things were also to be watched and roughly adjusted: the reading on an ammeter and the speed of a motor. These required some visual and auditory attention respectively. The psychologist, sitting near and directly in front of the reactor, watched his ability to control and keep up with the three tasks and, as the oxygen decreased, secured a progressive clinical impression of the reactor's efficiency at different levels without having any of the tasks provide a graphic or quantitative record.

It was the notion of developing some quantitative test which would be suitable for studying progressive asphyxiation as well as the desire to have a continuous reaction apparatus adapted for the measurement of a subject's efficiency under different conditions of nutrition, drugs, muscular work and fatigue which is responsible for the present development. The following conditions are desirable of combination in such a testing apparatus: (1) The result of a reactor's performance to be immediately available both during and at the close of the test, and this without any further measurement or computation; (2) the result to be a logical combination of the quickness and accuracy of the performance; (3) the apparatus to give an integrated value not only for the reaction time and reaction errors but also for the adequacy of the coördinated movements which normally follow a reaction, and for which the reaction is only the natural beginning; (4) the task to be a continuous one, fairly uniform in difficulty from minute to minute and yet so varied as to require the constant watching of the one being tested; (5) the amount of muscular work required of the subject to be relatively small and to admit of a sitting posture and not to be particularly fatiguing. The muscular coördination involved to be simple and logical in connection with the end to be attained; to be free from distraction, to necessitate quickness, accuracy, and judgment and place no upper limit on the possibility for improved performance; (6) the task in its time relations to be repeatable for standard tests on the same individual or different individuals and to be applicable to children and adults of varying grades of intelligence; and (7) the apparatus to be fairly independent of any laboratory condition, thus making it usable in all sorts of surroundings and capable of combination with other physiological or psychological measurements.

Description of the Pursuitmeter

(*a*) *The General Scheme*[1]

The task selected for testing purposes is that of continuously endeavoring to maintain the indicator of a wattmeter at a central zero position on its scale, by compensating with a variable resistance in one electrical circuit, changes in the amperage which are automatically induced in a second circuit through the action of an inductor device beyond the control of the subject. The arrangement is a combination of commercial electrical instruments adapted to the purpose by certain changes, principally in the winding of the field coils. The general wiring diagram is shown in Fig. 1. *A* is a wattmeter of horizontal edgewise type and zero center scale. This is viewed by the reactor and might be termed the stimulator. *B* and *C* are two test meters for integrating the errors of compensation, *B* accumulating those errors which the reactor allows to occur on the left-hand side of the zero position of *A; C* accumulating those errors on the right-hand side of the zero position. *D* is a curve drawing watt-

[1] For helpful suggestions in the early phase of planning, I am indebted to Dr. L. T. Troland. When spoken to concerning an electrical equipment for such purpose, he suggested the possible modification of a Wheatstone bridge employing an integrating meter in place of the telephone and requiring a subject to compensate with a variable resistance in one arm of the bridge variations which were automatically produced in the other. What has finally evolved is clearly related to that schematic idea.

To the kind coöperation of Mr. W. B. St. Clair, superintendent of the Standardizing Laboratory, General Electric Company, Lynn, Mass., is due the credit of providing a practical way by which common electrical instruments, all of which are produced by that company, could be so modified and combined as to attain the result which was sought. The willingness of Mr. A. L. Ellis, Mr. St. Clair and his assistant Mr. S. C. Hoare to advise and to have certain modifications of the electrical instruments made under their supervision is a coöperation, both personal and of the General Electric Company, which I cannot easily repay. I wish to acknowedge also the help of Mr. W. E. Collins, former mechanician at the Nutrition Laboratory and of my assistant, Mr. E. S. Mills, in the construction of the apparatus; the latter has especially aided me in collecting data for the Pursuitmeter test.

meter with zero center scale which draws in reduced amplitude the fluctuations of the needle of *A*.

In each of the four instruments, *A*, *B*, *C* and *D*, the field coils, which are, of course, stationary, were wound differentially, that is, with two wires in place of one, so that each instrument has four field binding posts. The number of turns and length of wire in both field circuits of an instrument were made the same and in the case of *B* and *C*, both instruments were made the same. When equal currents pass through the two windings, the two fields in an instrument oppose or buck each other and thus neutralize each other so

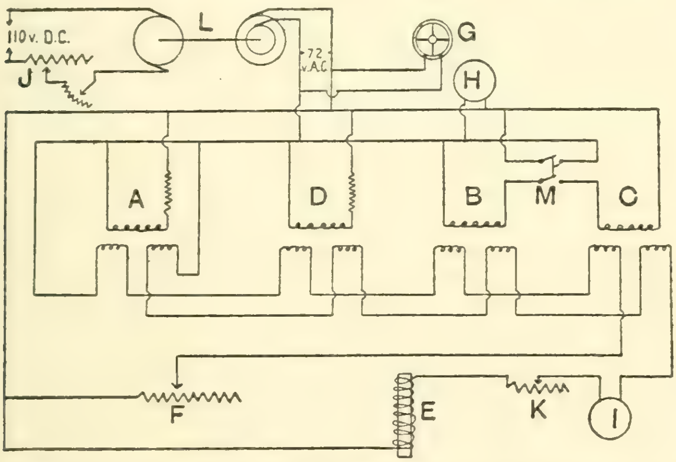


Fig. 1. General wiring diagram for the combination of electrical instruments which make up the Pursuitmeter. For lettering on figures, see text.

that no movement occurs in the armature of the instrument. The amperage in one circuit is disturbed by a combination of harmonic motions in the induction apparatus. *E* sets the task for the subject who seeks to compensate with the slide wire regulator, *F*. The disturber mechanism is operated by a motor, *G* with a worm gear reduction for driving the countershafts. *H* and *I* are voltmeter and ammeter respectively for standardizing the current in the disturbed circuit which may be adjusted by rheostats *J* and *K*, the former controlling the speed of the rotary converter *L*, which at the Nutrition

Laboratory was necessary to provide alternating current. As is well known, the moving systems of *A.C.* instruments and meters are very much lighter and more quick in their action than are those of the *D.C.* type. Hence this form was chosen as the more suitable for combination in such apparatus. The two wattmeters *A* and *D* are, of course, not dead beat. This is no special disadvantage. It only adds somewhat to the difficulty of the task.

The subject under test thus has not a number of discrete separate stimuli that are to be reacted to but a series of fluctuations which he is to parallel as nearly as possible. He watches the wattmeter, *A*, for his cue and does not know which way to move the regulator until the needle of *A* has started to leave the zero position, then having perceived the direction, he sees also something of the speed and governs his own compensation accordingly. The needle of *A* usually oscillates about its zero and the integrators *B* and *C* alternately turn up by small amounts and so accumulate the score. It is really a hare and hound game where you are to keep immediately at the rabbit's heel no matter what the direction of his jump or how quick or how far. There will always be loops where the two paths do not exactly coincide. These are the integrated areas which make up the score. They can never be entirely avoided but the object is to make them as small as possible.

The task outlined corresponds to that of holding voltage or steam pressure constant in a laboratory where instruments have to be standardized. It is somewhat like training a motion camera or a gun on a travelling object, steering a boat by the compass, trailing another automobile through crowded traffic and unfamiliar territory or driving on uneven car tracks. The same type of response is required in much industrial work. Much machinery necessitates not only a quick reaction but also a fairly definite coördinated movement following the reaction and this to be executed according to a fairly fixed pattern. There is thus reaction and pursuit movement for the purpose of compensating, directing, aiming, or otherwise tending these moving mechanisms. Undoubtedly it is possible to use industrial machines for measuring the individual's efficiency on the basis of his output, other conditions remaining suitable. Such mechanisms, however, are devised for manufacturing operations rather than to be used as psychological or physiological instruments in measuring the fluctuations of efficiency.

(*b*) *The Stimulus and Regulating Set*

To avoid distraction and for purposes of convenience in manipulating the subject and his surroundings, the stimulus wattmeter and the rheostat regulator are mounted together

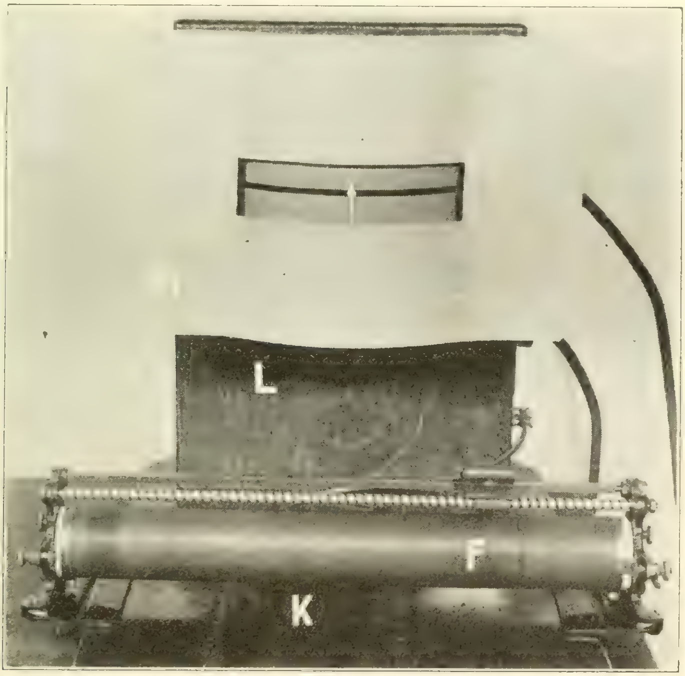


FIG. 2. Front view of the stimulus wattmeter and the regulator rheostat. By the manipulation of the rheostat a subject endeavors constantly to keep the wattmeter needle on the white zero line as illustrated.

in a unit which may be located according to desire in the same or another room from the other portion of the apparatus. The stimulus unit is shown in Figs. 2 and 3. On a suitable platform *K*, 36 x 72 cm. in dimension, is securely fastened a small cabinet *L*, 40 cm. high. To this cabinet the wattmeter, *A*, is attached as it would be to a switchboard. The accompanying resistances for the wattmeter are placed inside the cabinet, arrangements having been made for cooling ventilation.

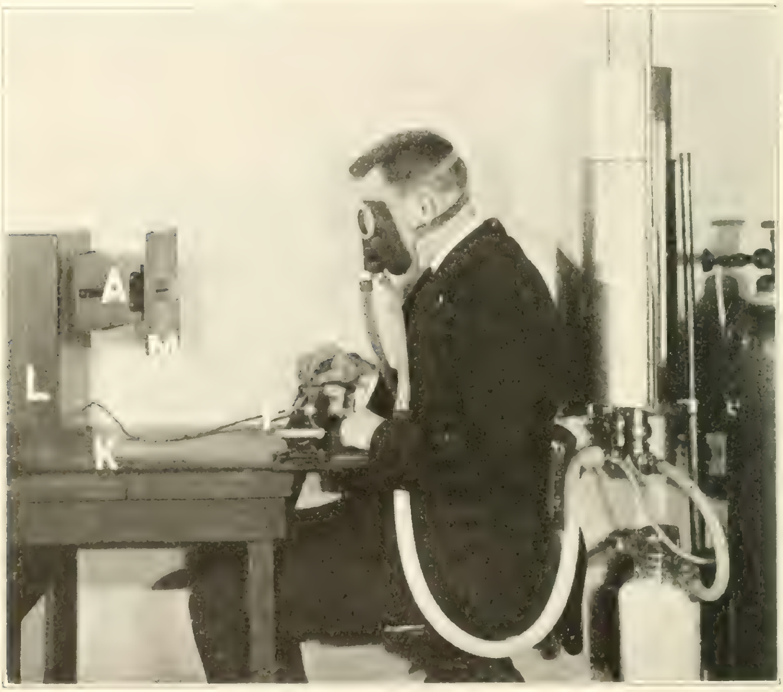


Fig. 3. Side view of the stimulus unit showing a subject being tested on the Pursuitmeter at the same time that his oxygen consumption is being measured on the "Benedict Portable" respiration apparatus.

Any task which requires continuous gazing at one object or area will cause vision to become at times somewhat blurred and after-image and contrast effects to be troublesome. This applies to the present apparatus as also to that of Seashore, McDougall and McComas. To make the visual conditions as satisfactory as possible, a screen *M*, 23 x 34 cm. in extent and faced with warm gray oatmeal paper was attached to the front of the wattmeter and the scale of the wattmeter behind the curved glass was faced with the same material. This has proved to be a comfortable background upon which to observe the movements of the diamond-shaped needle which is white as in the figure. The white line which extends across the scale is the zero position on which the reactor tries to hold the needle, the ideal position being that shown in Fig. 2. No other marks are placed on the scale. The reactor is instructed to keep the white diamond as nearly as possible exactly in register with the white mark. Provided the scale of the wattmeter is not brilliantly illuminated, fixation troubles are but slight. There is nothing to prevent the subject moving his eyes or his head to relieve tension. However, he can not look away far, or for an appreciable time and keep up with his task.

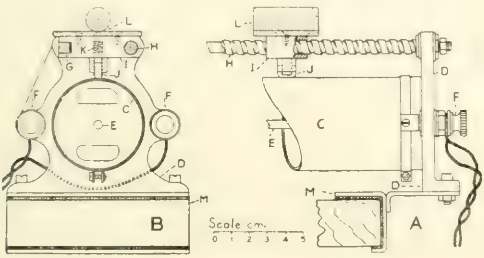


Fig. 4. Detail of the rheostat used as regulator, illustrating unusually rugged construction.

The compensation movement involves only the left or right hand according to the subject's preference and is direct, that is, he moves the rheostat slider in the direction that he wants the wattmeter needle to move. The coördination is a perfectly natural one. A six-year-old child can easily understand what is to be done and can do it after a fashion. The regulating rheostat is an important feature of such an apparatus. It should be of rugged construction, to avoid much wear and any loose parts or poor contacts and its slider should move with the least friction compatible with good contact on the resistance coil. These features were especially in mind in selecting and adapting the rheostat shown as *F*, Figs. 2 and

3, and in diagram, Fig. 4.[1] In this latter figure, *A* and *B* present side and end views respectively. The porcelain enameled tube *C*, is 16 inches long and wound with No. 25 Ideal resistance wire. The total resistance of the rheostat is about 400 ohms and the constant current capacity 1.2 amperes. The rugged construction may be seen from the heavy end pieces *DD*, the connecting rod *E*, the large binding posts, *FF*, and the two rods at the top, *G* and *H*. The former is square and the latter round and cut into a screw. The screw retains some oil and keeps the slide running smoothly. Upon these parallel rods *G* and *H* the slider, *I*, moves very freely and maintains a constant but very light pressure on the resistance wire. A shoe with a hard polished surface extends from the lower part of the slider and is kept in light contact by the action of a small coil spring, *K*. It is of prime importance that the slider shall not bind or stick or require anything but the slightest push of the hand to move it, otherwise the subject, who is trying to do careful compensating becomes exceedingly annoyed. The action is so easy that if the rheostat is tipped up vertically the slider will of its own weight move down to the bottom end. A large surface of fiber with a convenient portion, to be grasped in the fingers, is secured to the upper part of the slider. The rheostat is very conveniently and comfortably manipulated by the subject. Heating is never troublesome. In place of screwing or clamping the regulator to the baseboard which extends in front of the wattmeter, suitable guides, *M*, lined with cloth, have been secured to each end of the rheostat and these, passing one on either side of the baseboard serve not only to hold the rheostat firmly in place, but to make it adjustable as to its distance from the wattmeter. The subject in his preliminary adjustments can easily move the rheostat so as to best suit his own comfort.

(*c*) *The Integrating and Recording Instruments*

The two test meters for integrating the accuracy of compensation, together with the curve drawing meter, voltmeter,

[1] Made by H. E. Crane, Cypress St., Brookline, Mass.

ammeter and disturbing mechanism are built into a cabinet which is 80 cm. high by 70 cm. wide and 40 cm. deep. The upper part of this unit, showing the faces of the meters, may be seen in Fig. 5. The curve drawing meter is located on the top of the cabinet together with rheostat, *J*, which is in the 110 D.C. line and is for regulating the speed of the rotary converter and so adjusting the A.C. voltage shown in the voltmeter *H*. The push switch, *M*, for connecting the

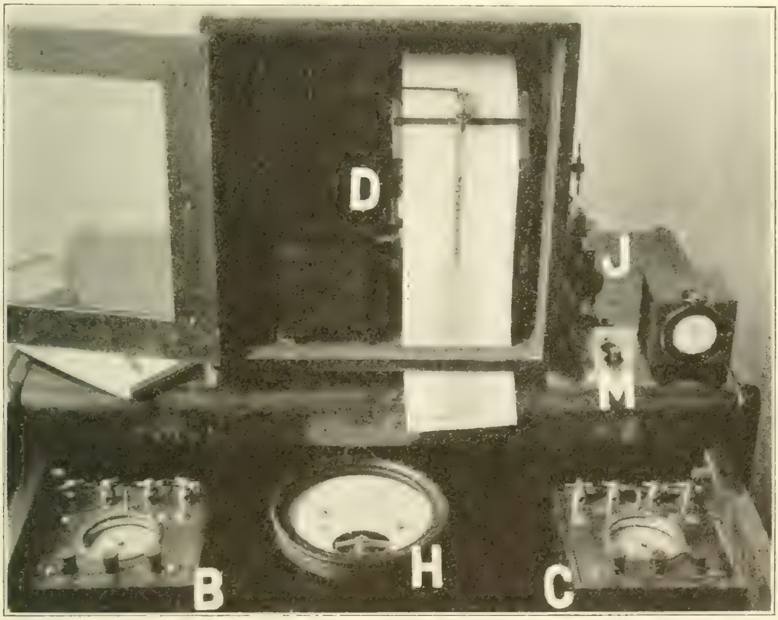


FIG. 5. Top view of the cabinet on which the integrating and recording instruments are mounted. This unit may be at any desired distance from the subject who is being tested.

potential sides of the integrating meters and so starting the test from zero position of the meters as well as stopping the test on the second is also on the top of the cabinet. The complete disturber mechanism is mounted inside the cabinet, the two integrating meters, *B* and *C*, are secured on the doors and the voltmeter is hinged from the top of the cabinet so as to come forward into a convenient position for observation by the experimenter. A glass plate lying across the top of the two doors provides a convenient desk and makes possible the observation and reading of the meter as desired. Without disturbing any part of the circuit, the doors may be closed for convenience in keeping out the dust, or when moving the apparatus.

The curve drawing instrument, *D*, is the most expensive part of the combination and it is not absolutely essential. It could obviously be used for a curve drawing apparatus and also to take the place of the stimulus apparatus. However, it is more sluggish in its movements, more distracting in its operation and also the subject could see the graphic record or some part of it. Furthermore, it is hard to watch and adjust to the zero position the rather clumsy ink pen which is a siphon arrangement. A final considération against using *D* as the stimulator is that its zero position is not very constant. The curve drawing meter is, however, quite useful in directly plotting the curve for the fluctuations caused by the disturber and the combination curve, made during the test, by the disturber and the reactor when the latter is trying to keep the current changes in the two circuits directly parallel. These records may be shown to the subject with a view to securing his better coöperation in further tests and it is thus unnecessary to make known his quantitative scores. This graphic record is not to be measured but indicates fairly well the consistency of the subject's work throughout the test period. Should he forget and releasing the rheostat slide, raise the reacting hand to his face, if it is necessary to cough or sneeze, or if he is distracted and looks about the room, the graphic record will show a loop in the curve which indicates clearly a lapse on the part of the subject. In standardizing the Pursuitmeter, one time with another, graphic records are useful and particularly in comparing the tasks when the movements of the induction apparatus are made simpler or more complex.

The integrating meters are known as portable test meters and are fitted with a zero set back arrangement. The dial is of white enamel, 6 cm. in diameter and divided into 100 divisions which are easily legible. The moving element is an aluminum disc on a vertical shaft to which a large hand is attached. This hand indicates directly each revolution of the shaft which may be easily read to the 100th part of a revolution. There are two small hands with corresponding dials located within the large dial. One of these pointers makes a revolution for every 10 revolutions of the large hand, the other for 100 revolutions. The meters as modified were rated for one ampere on 72 volts. This was also the rating for the two wattmeters.

It is apparent from the general wiring diagram in Fig. 1 that a common test meter with the field coils wound differentially would turn in one direction when the regulated circuit was too low in its amperage and in the other direction when this circuit had an amperage above that of the disturbed circuit. Starting with the meters at zero if the combined errors were as great on one side as on the other, the final reading would still be zero. Some delicate and effective arrangement must be provided which will allow the meter to turn in only one direction and so cause it to accumulate the errors in one direction, while a second meter performs a like service for the errors on the other side. Such a device is shown diagramatically in Fig. 6. On the armature shaft *A*, which is directly connected to the large hand shown on the dial of the integrating meter, was mounted a small brass wheel, *B*, 12 mm. in diameter and 1 mm. thick. On the periphery of this wheel 300 teeth have been accurately cut. A light pawl, *C*, is arranged to engage with these in such a way that the shaft of the armature can turn in only one direction. This pawl is carefully mounted in end-stone, ring-stone sapphire jewels; it has but slight back-lash, and is kept in delicate contact with the wheel by a hair spring with adjustable tension arrangements at *D*. The pawl and its bearings are mounted in a heavy block, *E*, which is rigidly connected to the frame of the test meter. By this wheel and pawl arrangement the integrated errors on one side of the zero position are accumulated and the sum for the two meters provides the total score. These meters have a zeroizing mechanism which consists of the usual interacting cams upon which pressure is brought by a small lever, *F*, and certain springs. The abuse of the delicately adjusted pawl, *C*, would be suicidal and is avoided by arranging a projection, L, from the plate, *M*, in such a position that when the plate is pushed forward to zeroize, the meter, *L*, engages with the counterweight of the pawl, *C*, and gently swings it out of contact with the wheel thus allowing free turning of the shaft, *A*, in either direction as necessitated by the position of cam, *J*, on the armature shaft. The spring, *N*, keeps *L* out of contact with *C* unless pressure is voluntarily exerted at *F*. The mechanism illustrated in Fig. 6, while rather delicate, in

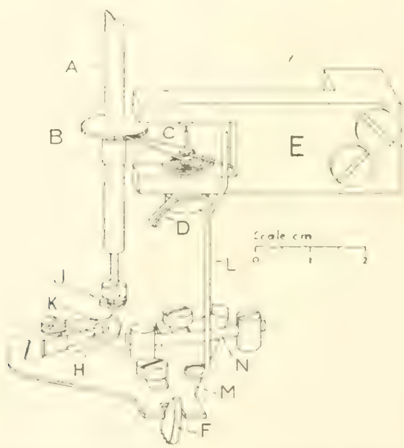


Fig. 6. Scheme whereby a pawl and fine toothed wheel arrangement makes it impossible for an integrating meter to turn in but one direction, hence it accumulates the errors.

reality remains in perfect and satisfactory operation for long periods at a time. As the scale on the dial of the test meter is graduated into 100 parts each such part corresponds to 3 teeth on the wheel, *B*. This degree of exactness in accumulating any small movements of the armature is wholly satisfactory as even a well practiced reactor will allow usual movements of from 2 to 8 divisions on the scale.

(d) *The Disturber Mechanism*

Of the many arrangements which could have been made to disturb one circuit and thus produce the task which will occasion the compensation on the part of the reactor a variable inductance was chosen as being suitable. It is relatively permanent, easy of operation, nearly free from friction resistance, and relatively independent of any changes from wear. Originally the disturber was a combination of three induction units, *i.e.*, three coils with as many soft iron cores. These were moved up and down by independent counter shafts run at different rates. Each unit could be made to function in different degrees since between each two layers of the winding on a spool, a tap was taken out, which made it possible to short circuit sections of the winding, so that each independent induction unit could be made to add its particular component to the complex of fluctuation in that proportion which was desired. Practically that arrangement was more complex than necessary. There was, furthermore, positive disadvantage in it since with three cores to be moved at the same time in and out of their respective coils, one or two were most of the time sufficiently within the coils to introduce a considerable amount of reactance in the circuit with the result that the amperage was usually low and varied through quite small changes and only at rare intervals reached its higher value. The form of disturber mechanism that has proved successful is shown diagramatically in Fig. 7. On the bottom of the cabinet, *A*, the spool, *D*, of the induction apparatus is mounted in such a position that the core, *C*, can move freely in a vertical direction in and out of the winding. The winding is of enameled copper wire (.042 inch diameter), ten layers deep and comprises 1,415 turns with a total resistance of 5 ohms. A tap is brought out between each two layers. The spool, *E*, is of brass and is 17.5 cm. long, inside length. Both tube and collars of the spool are split along one side to avoid heating from eddy currents. The core, *F*, is of soft iron wire firmly bound together and is tapered from top to bottom, the longest part being 18 cm. in length. This core is fitted into a fiber tube, *G*, which is somewhat longer than the iron core and moves very easily in the spool. The core contains about 800 g. of iron and complete weighs less than 1 Kg. By a cord passing over pulleys at *H* and *I* the core is connected to a counter weight, J, which weighs about 650 g. This serves to reduce the load on the driving motor, not shown in this diagram. The core, by the use of three countershafts, K, L, and M, is made to move vertically through a very complicated series of inharmonic motions or waves. Each countershaft is independent but all are driven by pulleys of different sizes on the common shaft, N, thus the speed of each counter-

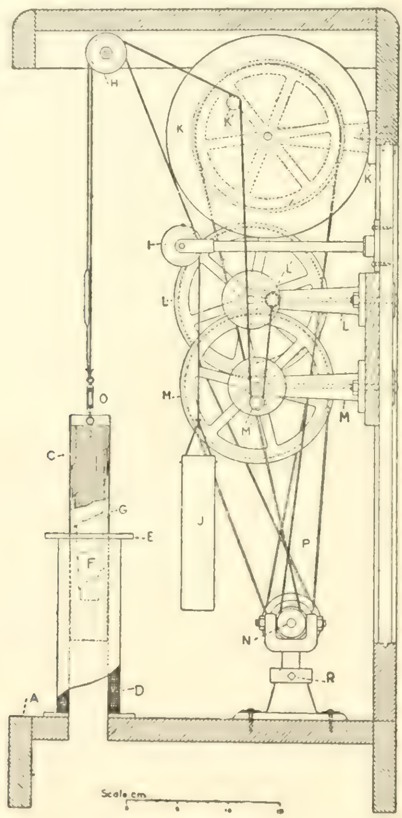


FIG. 7. Diagram of the disturber mechanism which is an induction apparatus the iron core of which is arranged to be carried through a complicated series of inharmonic motions.

shaft is different from that of the others as the driven pulleys above N are all 15 cm. in diameter. Each countershaft at its approximate end carries an eccentric which is at the same time a small pulley. (See K', L', and M'.) The cord through which the movements of these eccentrics are to be transmitted to the core of the induction apparatus is fastened around L', passes under M', over K', and over a second pulley at H, not shown in the diagram, and thence is connected to the turn buckle, O, which supports the core and allows for suitable adjustment so that when at its lowest point it will just correspond with the winding. The three eccentric pulleys, K', L', and M', as commonly used (of course, they can be otherwise located) cause the core to move vertically 95, 40, and 57 mm. respectively, a total of 192 mm. The three belts, P, which transmit the power to the three countershafts are quite positive in their drive. The speed is not rapid and the load is relatively small. The shaft hanger, N, is adjustable at R, so that a certain range of tightening may be provided. However, a small amount of slipping of the belts adds to the complexity of the current variations which the subject has to compensate and is of no special disadvantage.[1]

Manipulation and Characteristics of the Pursuitmeter

The arrangement of three eccentrics just described was the form of disturber arrived at empirically. Preliminary trials were made with a simple wave form caused by periodically moving the core of one induction apparatus, using only one eccentric, K', and giving it sufficiently wide amplitude to produce the total movement of the core. Such a wave form of simple harmonic character can be seen in Fig. 8, A, which is a tracing from the curve drawing wattmeter. Using a wave form of this character with a slow period, an efficient subject can almost succeed in perfect compensation. When the period is made more rapid he will not succeed in keeping the needle at zero every instant. However, the task is so uniform as to be monotonous and does not challenge his effort sufficiently. If three eccentrics as described above are used and these given rotation rates such as to cause their combination to result in a non-periodic series, a wave form

[1] It is recognized that it would be rather difficult to make a second disturber that would have exactly the same characteristics as this one.

something like Fig. 8, *B*, is produced. This series of fluctuations made entirely by the disturber has maximal and minimal amperages of 1.5 and 0.4 respectively. The amplitude of the largest deflections on the stimulus wattmeter, *A*, in Figs. 1, 2, and 3, are 40 to 42 mm. both right and left of the zero position. The curve drawing wattmeter plots these fluctuations just half size, hence in Fig. 8, the vertical lines of these sample records are 2 cm. apart.

From the plotted wave form of the curve drawing apparatus, the pattern for the task which has been used as standard requires approximately eight minutes to repeat itself. This agrees with what should be found since the

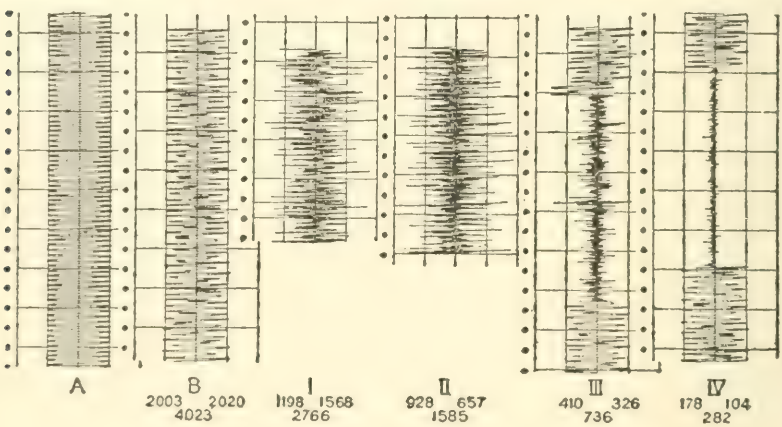


Fig. 8. Sample records made on the Pursuitmeter.

rotation speeds of the eccentrics, *K′*, *L′*, and *M′* are about 8.72, 4.03 and 5.40 seconds each which will divide nearly evenly into 480 seconds. The clock drives the paper for the curve drawing meter at the rate of one inch per minute but the periodicity of its escapement and small inaccuracies in the feed of the paper naturally change somewhat the curve which is plotted so that it seems never actually quite to repeat itself.

It is well within possibility that some subject should practically memorize the series of fluctuations if he had simply to watch them succeed each other. At least, it could be done if the subject set himself to it. It must be remembered, however, that by his manipulation of the regulator rheostat in compensating the fluctuations, he disturbs the whole series so far as the curve drawing wattmeter and stimulus wattmeter are concerned. Therefore, by the position of the stimulus needle in relation to its previous movements, he cannot anticipate what it is going to do next especially in reference to speed and amplitude of movement. He may anticipate that it will not move any further in a certain direction than the position it has taken judged by the position of the slider on the rheostat. There is thus definitely, the possibility of a kinesthetic memory since the hand in moving the regulator must go through a series of lateral movements which closely resemble and are about of the same amplitude as the series made by the core of the induction coil, provided that compensation is about adequate. The amplitude of the hand movements will vary from 4 or 5 mm. up to 19 cm. There are probably some subjects who, after long practise, will pay some attention to this kinesthetic imagery but it is much easier and more accurate to watch the deflections of the needle rather than to trust to this memory, and the latter will usually play but small rôle.

The stimulus needle starts and stops about 20 times a minute. In a large majority of cases, it starts to move in the opposite direction from which it has been moving. This the subject soon learns. However, he must guard against too great an anticipation, since occasionally it stops and continues again to move in the same direction or moves but slightly the opposite way. When it starts, its speed and amplitude of movements are two factors which he can hardly judge, but must wait for and meet as best he can at the instant. Speaking generally, the subject is required to react about every three seconds and in the interval between reactions to compensate a moving system as adequately as possible. The more successful subjects make a rather steady, smooth movement with the rheostat slider, jerky extreme movements cause added deflections due to the period of the instrument and confuse to some extent. If the needle is allowed to escape far either side of the zero position, recognition of the fact may cause something of a block in the reactor's work but is probably not quite such a disturbance as reaction errors with some other types of apparatus, since here it is a matter of degree rather than seeming to be definitely wrong unless the slider has been pushed in just the wrong direction and through a large amplitude. In this connection there is one feature against which the experimenter must caution his subjects. When the needle is quiet at a position that is fairly close to the zero line, some reactors let it rest and wait to see what is to happen next. Of course, the score tells the story and the plotted curve will show the larger errors, but the subject must be encouraged to keep actively at the job all the time he is being tested.

Certain very intelligent subjects, realizing that the integrating meters as well as the wattmeters, even though all A.C. instruments, have latency, have tried to beat the game by causing the needle to oscillate constantly across the zero line. The score has always been larger, *i.e.*, poorer, by this mode of compensation. While it would be desirable to have the armatures of these instruments move with the quickness and aperiodic characteristics of a string galvanometer, it is, of course, impossible and happily unnecessary. In adapting the instruments attention was given to making such changes as would reduce latency and over-shooting and since these factors are relatively constant they offer no special handicap to using the Pursuitmeter for comparative measurements, the purpose for which it was designed.

Another criticism may be raised which applies to the rheostat used as regulator. If a pointer is attached to the slider and a graphic tracing made of its movements while a skilled subject compensates a standard series of deflections as shown in Fig. 8, *B*, it will be found that from a neutral position the slider must travel much farther to the left than to the right. In fact from the position at which the slider remained constant during the time the tracing *B* was made 3.5 cm. of movement to the right will compensate the largest deflections to the left, while 15 cm. on the left will be required to balance the complementary deflections in the opposite direction. Such a rheostat is a linear arrangement and the electrical resistance represented by 1 cm. of its length is effective in changing the amperage of the circuit in proportion to the resistance already in the circuit. One could wish for strict proportionality in the size of the coördinated movements required by the test, practically subjects adjust themselves quickly to this discrepancy and their errors are approximately equal on each side. They have not to continually remind themselves that smaller movements are to be made on one side than on the other; each coördination is, so to say, felt out bit by bit without special regard to where or how far the slider is pushed.

Discussion of Results

Should a subject under test for a period of 5 min. allow a score to accumulate in each meter of 300 points, *i.e.*, three revolutions of the large hand, this would be substantially equivalent to his having allowed the needle to be five millimeters off zero for as much as 2 min. on either side of the zero line. The score is stated in terms of revolutions of the integrating meters rather than in terms of watthours or coulombs, since the integrating instruments are not measuring strictly the amperage which passes through them but only a function of the difference between the amperage of the two fields. Various arrangements of other instruments were tried in combination with the Pursuitmeter to secure a basis for giving the score in terms of electrical units. These arrangements seemed to add only complications and not to make the result more intelligible or trustworthy. It is reasonable to compare the score made by a reactor against that shown when the apparatus itself is operating without any intelligent direction. The slider of the regulator may be placed at such a position that the deflections are of equal amplitude on either side the zero line; with a maximum in the disturbed circuit of 1.5 amps. from 72 volts after a five minute continuous run, each integrating meter will show a score of very nearly twenty complete revolutions, *i.e.*, 2,000 points; the total score for both left and right will be 4,000.

Referring again to Fig. 8, *B*, a five minute portion of this tracing which was taken as a test like the one described, and is the standard check of the Pursuitmeter's proper working condition, shows a score on the left integrating meter of 2,003 and on the right 2,020, total 4,023. Record No. 1, Fig. 8, was made by a girl eight years old. She had never before seen the apparatus and, until just prior to the test, knew nothing of what she had to do. The graphic record indicates rather slow movements and that at times for several seconds the needle was not made to cross the zero line. Her score is 1,198 for the left hand side, 1,568 for the right, a total of 2,766, which as compared with the 4,023, shows that some intelligent compensation was being exercised. She may be said to have compensated to about the extent of 30 per cent. Record No. 2 is the first trial of a young physician who also had never seen or heard of the apparatus before. After a practise of less than a minute, he made the score 928 left, 657 right, a total 1,585. This is about 40 per cent. of the uncompensated score. He therefore compensated to the extent of about 60 per cent. Record No. 3 was by an assistant who had had considerable practise with the apparatus throughout the period of its development and understood very well its mechanical and electrical features. His total score of 736 is only a little over 18 per cent. of the unbalanced score. About midway in this record, there is a large deflection seen on either side the zero position, coincident with which the subject coughed. Record No. 4 is by the most skillful subject who thus far has been tested. It indicates quick reaction and that each compensation is a very smooth and controlled movement, not just the pushing of the slider to the opposite end of the rheostat. The total score of 282 represents that compensation has been adequate to the extent of about 93 per cent. of the total value. If a subject is approaching too near a zero score, the reasonable thing to do is to increase the complexity or speed of the disturber fluctuations. It is desirable, however, for long periods of experimentation on the same individual and where different individuals are to be compared in reference to the influence of some outside factor, if the task can be made difficult enough to start with so that it will satisfy all requirements throughout the whole series without at the same time being so hard as to wholly discourage the unpracticed and less efficient subjects.

As illustrative scores for average adult performance on the Pursuitmeter some data are given in Table 1. These records were taken on several members of the laboratory staff, each person making five successive 5 min. trials per day. The average score for each subject on each day is shown in the table. All of these subjects were entirely new at the task. *Bw* and *Dw* were women and demonstrate marked and very consistent improvement. The men *C* and *E* also improved

TABLE I

AVERAGE ADULT PERFORMANCE ON THE PURSUITMETER. FIVE SUCCESSIVE 5 MIN. TRIALS WERE MADE BY EACH SUBJECT ON EACH DAY. THE VALUES ENTERED ARE THE AVERAGES. *Bw* AND *Dw* WERE WOMEN

Subject	Successive Days of Practice				
	1	2	3	4	5
Bw	1,330	945	872	774	772
C	1,020	1,061	839	919	812
Dw	1,621	1,372	1,241	1,094	1,073
E	1,714	1,426	1,468	1,677	1,392
F	2,423	1,759	1,559	1,412	1,276
G	2,501	2,055	1,656	1,429	1,378
Av.	1,768	1,436	1,272	1,217	1,117

by practise, as indeed any reactor should do, but they both were quite erratic and did poorly on the fourth day. Subjects *F* and *G* at the beginning did little better than the 8-year-old girl, but percentage-wise they improved more than the others excepting *Bw*. In general adults make a score of about 1,800 on the very first trial. The average for the group shows 1,768 for the first day of five trials and some improvement for each day thereafter, with an average score of 1,117 on the fifth day, *i.e.*, after twenty-five 5-min. periods of practice. In the case of four subjects, *A*, *C*, *Dw* and *G*, the results for all the practice periods have been plotted individually in Fig. 9. The scores are quite regular in every case. *A* made but little change during these 25 trials having had some earlier practice. The others show consistent practice effects. Between trials Nos. 25 and 26 there was a break of one month's vacation. *Dw* shows up definitely poorer during this latter period which correlates with her poorer physical condition experienced then. The other three continue their improvement in the test.

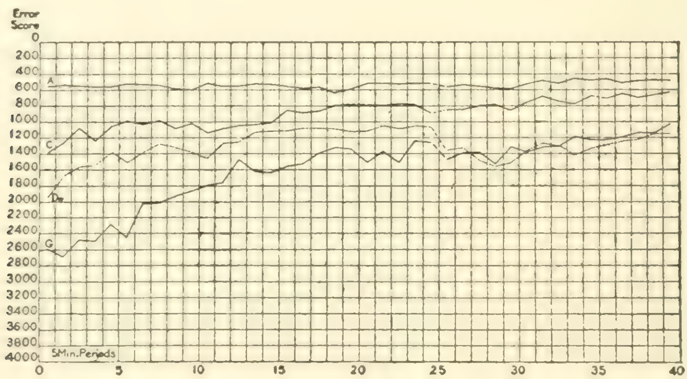


Fig. 9. Individual practice curves indicating the acquisition of skill in the Pursuitmeter Test. Between trials No. 25 and 26 a vacation period of one month intervened.

The Pursuitmeter has proved a convenient and sensitive test of neuro-muscular efficiency. It is, among other uses, quite applicable to the investigation of problems of nutrition when some superimposed factor is brought to bear on the human organism. Such use of the measurement may be illustrated by a bit of alcohol data from a research now in progress. The data are not given here as a contribution to alcohol literature; being on one reactor they must not be generalized. The presentation here is only to show illustrative data obtained with the apparatus. Following a light lunch the subject went through a series of measurements, 8 tests in all, which included the Pursuitmeter test. The series was repeated and after the second period he drank 1 liter of water or 1 liter of water in which 27.5 grams of ethyl alcohol had been diluted. The quantity of liquid and temperature were not varied. Fifteen minutes were quite sufficient for drinking the liquid, after which the series of measurements were repeated three more times, *i.e.*, periods 3, 4, and 5 of the day. Table II, which is typical of several which might be given,

Table II

Data Illustrating the Convenience and Sensitivity of the Pursuitmeter Test as a Measure of Changes in Neuro-muscular Efficiency as Produced by a Nutritional Factor

Date	Successive Half-hour Periods					
	1	2	Drink	3	4	5
1919						
Nov. 5	460	436	1 liter water	456	450	477
Nov. 8	407	370	" "	395	386	381
Nov. 19	365	357	" "	339	335	399
Nov. 21	426	378	" "	342	362	382
Nov. 22	354	341	" "	346	331	317
Av.	402	376		376	373	391
	389			−13	−16	+2
Nov. 6	436	424	1 liter 2.75% alc.	447	449	467
Nov. 7	417	378	" "	380	442	446
Nov. 10	390	374	" "	393	419	424
Nov. 17	398	387	" "	423	428	445
Nov. 18	367	320	" "	416	398	417
Av.	402	377		412	427	440
	389			+23	+38	+51
				36	54	49
				389	389	389
Av. loss, 11.9 per cent				9.3	13.9	12.6

shows data for five normal days and for five days on which alcohol was used. The values entered are the sum of the two meter readings for a 5 min. test period. The average score for the first periods of normal days is 402; the average score for the first periods of the five alcohol days is 402. Likewise the averages for the second periods for the two groups of days are 376 and 377 respectively. It is simply a matter of chance that these preliminary values are so identical. It must be recalled in interpreting these data that the smaller scores represent the better efficiency. Following the drinking of the water the subject was able to do just about the same as he did at the second period, or if we average the first and second periods to get his preliminary performance, we find that in general, he did better by 13, 16, and poorer by 2 points in periods 3, 4, and 5 after the water. On the other hand, following the alcohol, the compensation was poorer on the average then the subject had done in the preliminary periods. There is only one instance (third period, November 7, 380 as compared to 417) when the subject did as well or better after the alcohol than he had done in the poorer of his preliminary periods. The average decrements for the alcohol days as against the preliminary score, 389, are 23, 38, and 51. The differences between these and the performance when only water was taken show for periods 3, 4, and 5 losses of 36, 54, and 49 points on the basis of 389, thus corresponding to 9.3, 13.9, and 12.6 per cent., with an average loss for the three periods of about 12 per cent. The measurements represented by periods 3, 4, and 5 came approximately 15, 45, and 75 minutes after the ingestion of the alcohol. This table was chosen for illustration because practice was still producing considerable change in the scores. Although there was improvement in the subject's scores between November 5 and 22, still the data within any one day are uniformly consistent and each succeeding pair of days, if taken separately, show practically the same contrast in efficiency when measured by this continuous reaction-coördination apparatus.

Summary

1. For detecting the influence of superimposed nutritional factors, fatigue, industrial conditions and the like, it is urged that not only the reactions but the eye-hand coördinations that naturally follow reactions should be measured.

2. An apparatus called a Pursuitmeter is described in detail. It is an arrangement of common electrical instruments which tests the subject's ability to balance two electrical circuits. The task is continuous and requires constant attention calling for both quickness and accuracy. The amplitude and duration of all errors are integrated electrically and accumulated by two meters which may be read directly during or at the end of the test.

3. The measurement of children and adults may be directly compared. It is shown that practice results in quite regular improvement and that the apparatus provides a very convenient and at the same time sensitive indicator for neuromuscular efficiency.

REFERENCES

1. MILES. *Psychol. Rev.*, 1920, **27**, p. 361.
2. SEASHORE. Univ. Iowa, Studies in Psychol., Vol. 3; Bull. State Univ. Iowa, N. S., 1902, No. 49, pp. 1–17.
3. McDOUGALL. *British Jour. Psychol.*, 1904–05, **1**, p. 435.
4. McDOUGALL AND SMITH. British Medical Research Council, Special Report Series, No. 56, 1920, pp. 1–34.
5. BURT. *British Jour. Psychol.*, 1909–10, **3**, p. 153.
6. SMITH. *British Jour. Psychol.*, 1916, **8**, p. 329.
7. McCOMAS. *Jour. Exper. Psychol.*, 1917, **2**, p. 171.
8. DUNLAP. Air Service Medical, Washington, 1919, p. 300 ff.

FURTHER TECHNIQUE FOR INSPIRATION-EXPIRATION RATIOS

BY HAROLD E. BURTT

Ohio State University

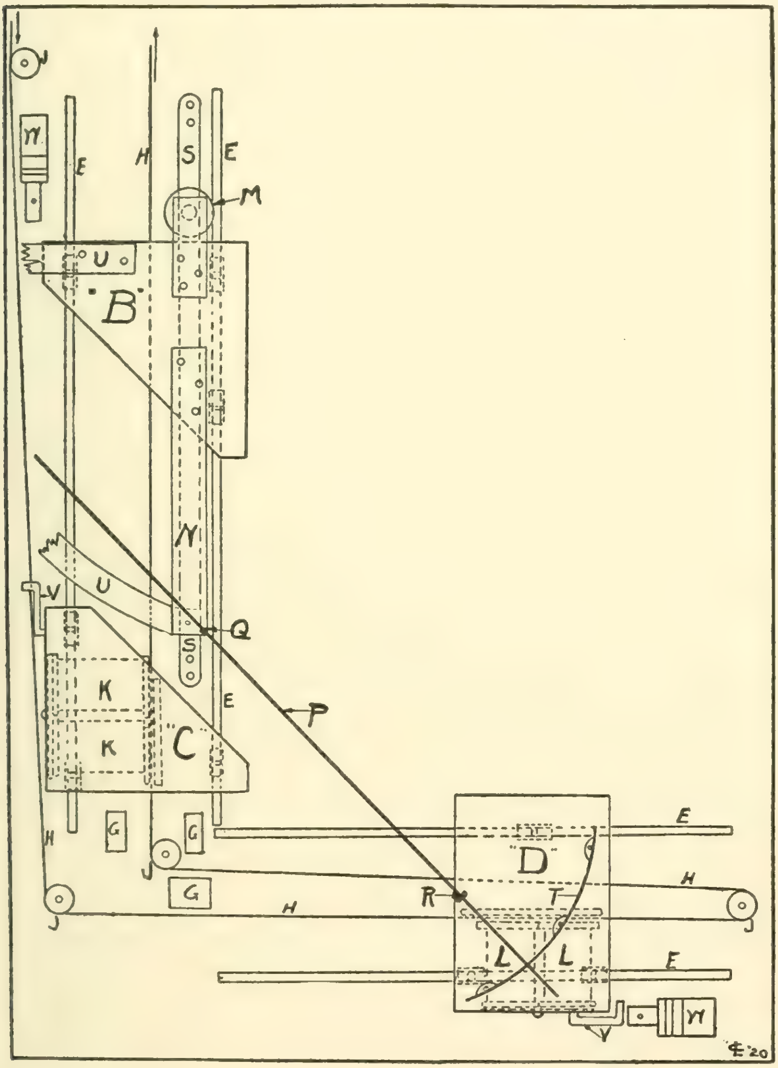
A device has been previously described for closing one electric circuit during inspiration and another circuit during expiration.[1] It consisted essentially of a one-way contact attached to a tambour operated by the usual pneumograph. This contact was in series with the magnet of a double relay and the two relay circuits could be used to trace two time lines (interrupted 5 or 10 times a second) on a kymograph or to operate two electric counters each of which recorded 5 or 10 units a second while the circuit through it was closed. In either instance the process of obtaining the I/E (inspiration-expiration) ratio was tedious. It was necessary to measure time lines or count them if they were in series with an interrupter, or to subtract the successive readings of the counters. The ratio was then computed on a slide-rule for each breath. The device described below is connected to the double relay, supplanting the kymograph or counters, and at the conclusion of each breath indicates on a scale the I/E ratio for that breath. These figures can be recorded during the experiment.

The apparatus is based on the fact that the ratio between the legs of a right triangle is the cotangent of one of the acute angles. If two members move from the right angle along the legs at constant speed one during inspiration and the other during expiration and if a straight line connects them at the end of a breath, the I/E ratio can be read in terms of the angle thus formed.

Three light wooden cars B, C, and D (see figure) each having three wheels move on two tracks E situated at right angles. The tracks are in the same plane but the whole apparatus tilted slightly so that gravity brings C and D

[1] Burtt, H. E., 'A Pneumograph for Inspiration Expiration Ratios,' *Psychological Bulletin*, 1918, *15*, 325-328.

against the stops *G* and brings *B* against *C*. The stops and the contiguous edges of *B* and *C* are padded to reduce the noise. An endless belt *H* (two lengths of ordinary string



twisted together) is driven at practically constant speed by a motor and worm reducer over the pulleys *J*. Cars *C* and *D* have electromagnets *K* and *L* mounted beneath them with armatures suspended from the under sides of the cars. The endless belt passes between the armatures and the poles of the magnets in such a way that if the belt is moving in the direction of the arrow the sections of it passing over the poles of the magnets are moving in a direction away from the stops toward the outermost ends of the tracks. When the current passes through the magnets the armatures clamp the cars to the belt and they move in the corresponding direction. A magnet *M* is mounted vertically on the front of *B* with the bottom end of its core very close to a strip of iron *S* of adjustable height and extending under the entire course of *M*.

Car *B* is pushed along by *C*. It carries a brass member *N* which projects over *C* when the two cars are together. A very light lever *P* is pivoted on this brass member at *Q* and passes through a guide *R* projecting from the top of Car *D*. This car also carries a curved rest *T* for this end of the lever which is the heavier.

When the apparatus is at rest *Q* and *R* are as close together as possible (theoretically they should coincide but a slight amount of play is mechanically necessary) with the lever parallell to the *BC* track. The pneumograph previously described operates the magnet of a double relay. During inspiration the current from the relay passes through *K* and the two cars *B* and *C* and hence the pivot *Q* travel at constant speed up the incline with the lever *P* remaining in the same direction but lengthening its portion between *Q* and *R* directly in proportion to the length of the inspiration. During the following expiration the current is broken through *K* but is made by the other pole of the relay through *L* and *M* in parallel. Car *B* is immediately clamped to the rail *S*, while *D* and hence the guide *R* move at right angles to the direction taken by *Q*. The distance traversed by *R* is directly proportioned to the expiration time. The lever forms a gradually lengthening hypotenuse of a triangle and the cotangent of the angle formed between the lever and a line through *Q* parallel to the track is a direct index of the inspiration expiration ratio. A cardboard scale mounted on *B* (only a portion of the frame *U* supporting the scale is shown) is traversed by the shorter arm of the lever and graduated to read directly the cotangent of the angle.

To avoid injury to the apparatus by a very long breath the cars *C* and *D* carry projecting arms *V* which will, as the car approaches the end of the track, strike the circuit breakers *W*. These consist each of a piece of brass hung vertically from a hinge and resting under gravity against a spring brass contact. These breakers are in series with the respective magnets. The arms *V* push the swinging members away from the contacts thus breaking the magnet circuits, releasing the armatures and allowing the belt to slip through without pulling the cars farther. If a breath is unusually long and either circuit breaker operates the car can be seen to vibrate back and forth slightly in its position near the end of the track and that breath can be omitted. During expiration the car *C* moves back to the zero position under the influence of gravity and is ready for the next inspiration, but car *B* remains in its position until the ratio is indicated by the lever. At the end of expiration *B* and *D* start back and the reading must be caught at that moment. *C* starts again with the following inspiration (it has ample time to return during the preceding expiration). *C* meets *B* coming back and carries it along until the end of inspiration and *D* has meanwhile returned to the starting point. The use of the additional car *B* obviates the necessity of having two identical pieces of apparatus working alternately.

The present model of the apparatus is about 14 x 18 inches outside dimensions. With this size the scale can be read easily to 0.1 and for smaller ratios to 0.05. With a little practice one can catch the reading at its maximum point as the lever starts to swing back to the initial position. The speed of the belt may be adjusted for slow or rapid breathing by attaching it to a different portion of the cone reducer connected with the worm. The magnets are wound with No. 36 wire and each run in series with a 25-watt lamp.

The time saved by the device is considerable. In some experiments now in progress, there is a clear saving of three hours in reading kymograph records for every hour of experimentation. The reliability of the results is also increased because the experimenter can note any unusual conditions such as coughing or premature speaking and omit those breaths from the record. Needless to say the technique is applicable only where the entire interest is in the I/E ratio and not in the length or amplitude of the breaths.

COMPLEX SIGNS IN DIAGNOSTIC FREE ASSOCIATION[1]

BY CLARK L. HULL AND L. S. LUGOFF

University of Wisconsin

Introduction

In 1904 Jung, Riklin, Wehrlin and other members of the Bleuler-Freudian school of psychiatry, began publishing in the *Journal für Psychologie und Neurologie*, an extremely significant series of studies on free association as a means of diagnosis. Some of these studies are somewhat theoretical and *a priori* in character, but many of them are thoroughly scientific and quantitative in their nature. In them two things stand out: (1) Jung's classification and analysis of the kinds of responses given by various types of normal and pathological individuals and (2) his enunciation of various signs (Komplexmerkale) by which he believed emotional complexes to be indicated. It is with the latter aspect of the free association method that the present investigation is concerned.

Jung seems to have noticed these signs first as a result of general observation of the reactions of the patients in his clinic while giving the free association tests. He appears to have noted marked coincidences between certain types of reaction and the presence of emotional complexes as revealed chiefly by introspections obtained from the patients. These introspections were secured by a special technique now well known by the name of psychoanalysis. Unfortunately Jung was content for the most part to base his belief in the various signs merely upon such general observations. With an exception to be noted later, he made little attempt to secure any exact or objective evidence as to the truth of these important assumptions. Yet his statements seem to have

[1] This is the first of a series of Wisconsin studies in free association.

been accepted with little question by psychologists and to have found extensive application in practical diagnosis.[1]

The unsatisfactory state of this field in mental pathology is shown by the fact that there are innumerable points concerning complex signs upon which we have no decisive evidence whatever. Indeed many of the questions seem never to have been raised at all, despite their obvious clinical importance. For example, do *all* of the alleged indicators really indicate emotional complexes? And of those indicators which do, how strong are the strongest and which are the weakest? Does long reaction-time really deserve the prominence usually given it?[2] And what is the diagnostic status of extremely short reaction-time?[3] Moreover, if long reaction-times really indicate complexes, exactly how long must they be before doing so? Is the transition in potency from non-critical to critical times, gradual or abrupt? And how does diagnostic potency above the critical point, vary with further lengthening of the reaction-time? Or again, does the significance of reaction-time lie in its absolute or its relative length or in some kind of joint relation of the two?

Or we may consider a single detail like the alleged perseveration tendency of which Jung makes so much.[4] Does it really exist apart from accidental successions of long reaction-times, say, and if so how strong is it? Does it exist in the case of all the complex signs or only for some? Is it equally strong for all signs where it operates at all, or does it vary, say, according to their respective diagnostic strengths? How long and how far in the series does it persist, and is the curve of its diminution one of positive or of negative acceleration? May perseveration take place in terms of a complex sign other than that of the critical (initiating) reaction, and if so, relatively how strong are the various possible combinations of critical and post-critical signs? Do two or more

[1] Brill, A. A., 'Psychanalysis,' Philadelphia, 1917, p. 140 ff.

[2] Jung, C. G., *J. für Psychologie und Neurologie*, IX., p. 189; also Jones, Ernest, 'Papers on Psycho-Analysis,' New York, 1919, p. 406.

[3] Dooley, Lucile, 'A Study in Correlation of Normal Complexes by Means of the Association Method,' *Am. J. Psychol.*, 1916, XXVII., p. 132.

[4] Jung, *op. cit.*, p. 189.

complex signs appearing in a single critical reaction, tend to mutual inhibition or facilitation of their normal perseveration tendencies, and how does this action vary with different combinations of signs and diagnostic potencies? It was with a view to calling attention to numerous problems of a similar nature and to contributing some suggestions toward the solution of a few of them in so far as they apply to normal subjects, that the present investigation was undertaken.

The data consists of 10,000 free association reactions obtained from 100 subjects, 50 men and 50 women. The subjects came from the various middle and upper levels of population from the cities of Chicago, Duluth and Madison. All possessed a grammar school education, the majority had completed the high school and a few were attending the University of Wisconsin. The experiments were in all cases conducted in a quiet room free from distractions. The experimenter was the same throughout.

The instructions given the subject were as follows: "I am going through a list of words. After every word I say, I want you to respond with the first word that comes to your mind as fast as you can, and no matter what the word is. It is as if I say the word 'table,' and the first word that comes to your mind is 'chair,' you say 'chair.' Do you understand? We'll try one word to see if you have got the idea. I'll say the word and you say the first word that comes to your mind as fast as you can and no matter what it is. All ready now—'Snow!'" If the subject then understood what was wanted of him he was told, "Now we'll run through the list." If he did not understand, other practice words were given until he showed that he had a clear idea of what he was to do. It was seldom necessary to give more than one practice word. He was then given the 100 words of Jung's well-known association list in the order given on p. 116.[1] The time required for the reactions was measured by a fifth-second stop watch and

[1] Jung, C. G., 'The Association Method,' *Am. J. Psychol.*, XXI., 1910. The present list deviated slightly from that of Jung in a few places. The most of the infinitives of the original have been changed to gerunds as being under the circumstances more nearly the psychological equivalent in English of the German originals.

recorded as made. Various complex signs not shown by the verbal response word and the time, such as repeating the stimulus word, laughing, etc., were recorded by a convenient set of arbitrary signs.

After all the words had been reacted to, the subject was instructed as follows: "I am going through the list of words once more and I want you to try to give me the same words you gave me before. If you remember or think you remember, say it. If you don't, say 'No.' You can take as long a time as you want, to answer." Whenever any peculiar or unusual reaction was made, subject was quietly asked to explain it before proceeding with the experiment. At the conclusion, elaborate notes as to any peculiarities of the behavior of the subject were made on the back of the recording blank. Lastly each record was gone over while fresh in experimenter's memory, and scored with great care by means of a system of arbitrary signs. One arbitrary sign was placed opposite the stimulus word for each complex indicator found connected with that word.

The complex signs investigated in the present study are as follows:

1. Long reaction-time (over 13 fifths of a second).
2. Inability to make any response whatever.
3. Extremely short reaction-time.
4. Repetition of the stimulus word itself.
5. Assimilation (apparent misunderstanding) of the stimulus word.
6. Defective reproduction of original reaction at second presentation of the stimulus word.
7. Response with the same reaction word at two or more different stimulus words.
8. Strange or apparently senseless reactions.
9. Perseveration.

II.

The Relative Power of Various Words to Evoke Complex Signs

At the outset, it is a matter of some interest to know with as much precision as possible the characteristics of the series of words used as stimuli. The list here employed is the result of years of careful trial by Jung and it is supposed to contain special properties.[1] Are some of the words very potent in revealing emotional complexes and the rest neutral, or is power widely and about equally distributed? Are powerful words placed in groups or at intervals or do they appear in the list according to chance? Are some words specially significant for one sex and others for the other sex, or are there no sex differences in this respect.

FIG. Showing the distribution of power of the respective words of Jung's list to evoke complex signs.

In an attempt to find at least an approximate answer to some of the above questions, it was assumed provisionally that all of the alleged complex signs here investigated really indicate emotional complexes; that all indicators are equally potent in this function; and that the greater the total number of indicators yielded by a particular word in the records of a large number of subjects, the more powerful that word is in revealing emotional complexes. The indicators were therefore tabulated by words separately for the fifty men and the

[1] Brill, *op. cit.*, p. 140.

Table I

Showing the Number of Complex Signs Yielded by the 100 Subjects for Each of the Words of the Jung List

	Men	Women	Both		Men	Women	Both
1. head	22	23	45	51. frog	23	27	50
2. green	34	36	70	52. to part	59	46	105
3. water	36	41	77	53. hunger	25	27	52
4. sing	34	26	60	54. white	36	40	76
5. dead	30	32	62	55. child	38	37	75
6. long	18	19	37	56. caring for	75	53	128
7. ship	34	23	57	57. pencil	17	20	37
8. pray	31	31	62	58. sad	38	38	76
9. window	26	24	50	59. plum	24	30	54
10. friendly	54	64	118	60. marry	41	54	95
11. cooking	31	24	55	61. house	25	18	43
12. ask	30	37	67	62. dear	47	50	97
13. cold	29	30	59	63. glass	32	42	74
14. stern	47	52	99	64. quarreling	40	37	77
15. dance	30	28	58	65. fur	15	25	40
16. village	19	13	32	66. big	23	26	49
17. lake	28	27	55	67. carrot	31	30	61
18. sick	23	29	52	68. painting	27	18	45
19. pride	66	49	115	69. parting	54	58	112
20. bake	34	40	74	70. old	20	28	48
21. ink	27	34	61	71. flower	32	36	68
22. angry	52	53	105	72. beating	57	54	111
23. needle	16	16	32	73. box	40	28	68
24. swim	39	42	81	74. wild	28	26	54
25. voyage	24	26	50	75. family	27	35	62
26. blue	40	46	86	76. wishing	32	38	70
27. lamp	22	22	44	77. cow	14	17	31
28. to sin	80	77	156	78. friend	46	52	98
29. bread	39	26	65	79. luck	30	32	62
30. rich	20	26	46	80. lie	37	36	73
31. tree	16	17	33	81. conduct	42	46	88
32. prick	41	46	87	82. narrow	13	23	36
33. pity	58	47	105	83. brother	15	11	26
34. yellow	49	43	92	84. fearing	55	62	117
35. mountain	16	12	28	85. stork	16	24	40
36. to die	37	36	73	86. false	43	43	86
37. salt	28	29	57	87. anxiety	45	33	78
38. new	23	24	47	88. kissing	30	41	71
39. custom	44	42	86	89. bride	24	19	43
40. supplicate	63	55	118	90. pure	27	26	53
41. money	22	26	48	91. door	26	24	50
42. foolish	29	33	62	92. choosing	30	46	76
43. pamphlet	26	29	55	93. hay	12	13	25
44. despise	37	19	56	94. contents	34	36	70
45. finger	13	12	25	95. ridicule	37	33	70
46. expensive	36	34	70	96. sleeping	20	11	31
47. bird	27	19	46	97. month	35	10	15
48. fall	27	49	76	98. nice	47	46	93
49. book	26	28	54	99. woman	35	26	61
50. unjust	41	32	73	100. abusing	50	48	98

fifty women, totaled, and these totals combined for the final determination of the potency of the individual words. These results appear in Table I.

A brief glance at this table shows that power of evoking complex indicators is very widely distributed. Not a single word in the list fails to show its quota of signs. That there is no break or even a considerable tendency to cleavage between strong and weak words is demonstrated very clearly by the curve of Fig. 1, which shows the relative strength of the 100 words arranged in the order of potency from the strongest to the weakest. Indeed the almost perfect ogive form of this curve shows that the frequency distribution of words in this respect is almost exactly Gaussian or according to pure chance. The interested reader may easily verify this by plotting a frequency polygon of the data in Table I.

As to the arrangements of the weak and strong words in groups or at intervals, relatively little can be discovered even by a careful inspection of Table I. Fortunately there is a method by which any given tendency to sequence of weak and strong words may be detected and its strength determined. For example, in Table I., of all the words which come immediately after strong words, 20 are themselves strong and 32 are weak; and of all words following directly after weak words, 32 are themselves strong and 15 are weak. Clearly weak words tend to follow strong words and strong words to follow weak. By a method to be described later (p. 125), the strength of this tendency to alternation may be determined with precision. In the above case it is $-.45$. This is about half way between a purely chance sequence and a perfectly uniform tendency to alternation of weak and strong. A similar computation for words following next but one, yields a coefficient of $+.25$. This is about a quarter of a tendency for both weak and strong words to be followed by one of their *own* kind, after an interval of one word.

In all, nine such coefficients were computed, determining the nature and the extent of the arrangement of the words for various distances following the respective weak and strong words of the list:

1st. following, − .45	4th. following, + .06	12th. following, + .14
2d. following, + .25	5th. following, − .41	25th. following, − .31
3d. following, − .31	6th. following, + .35	50th. following, + .29

The above figures show that Jung not only tried to have a weak word follow a strong one whenever a strong one appeared on the list, but the uniform alternation of plus and minus signs in the above table shows that the alternation was a definite program extending as far as 50 words distant. No doubt the coefficients would have been larger had it not been for the influence of the translation of the words from the German, and certain changes introduced into the list by the present writers. The distinctly smaller size of the plus coefficients is suggestive though its meaning is not quite clear.

The character of the words bearing the heaviest burden of complex signs next engages our attention. The Freudian literature leads us to expect these words to be mostly such as would have obvious connections with powerful emotions, particularly those associated with the reproductive instinct and such as are likely to be strongly repressed. Upon the whole this expectation is realized though with important qualifications. The 20 strongest words follow, together with the total number of indicators yielded by each:

to sin	156	to part	105
caring for	128	stern	99
friendly	118	abusing	98
supplicate	118	friend	98
fearing	117	dear	97
pride	115	marry	95
parting	112	nice	93
beating	111	yellow	92
angry	105	conduct	88
pity	105	prick	87

Perhaps half of the twenty words have a significance for erotic complexes, if interpreted in a broad sense. An even more striking tendency is the large proportion of words touching emotions of a distinctly painful coloring such as anger, fear and sorrow. This is particularly interesting in view of the recent tendency to assign a larger rôle to these latter emotions as a source of mental disorder. But of the strongly emotional nature of these words there can be little doubt. Perhaps this may be most clearly shown by comparing them with the ten weakest words:

tree	33	mountain	28
needle	32	brother	26
village	32	hay	25
sleeping	31	finger	25
cow	31	month	15

The qualitative contrast between the two sets of words is marked to say the least.[1]

Upon the whole the men and the women show about the same number of signs on the various words. The four words *pray*, *lamp*, *needle* and *sad* (Table I.) have exactly the same number of indicators for each sex, while 80 of the remaining 96 words vary in their difference from one to ten points, with an average of four. *To sin*, which yields the largest number of indicators of all the words for both sexes alike, shows a difference of only four in the two scores. The men yield 80 indicators while the women yield 76.

A few words, however, differ greatly in the number of signs yielded by the men and the women. The words *fall*, *caring for*, *despise*, *pride* and *choosing* show this characteristic in such a marked degree that it can hardly be due to chance. The word *fall* for example, yields a difference of 22 points with the higher score being given by the women. This is easy to understand in the light of the common use of the word *fall* in a moral sense. The respective scores of 27 and 49 may perhaps be considered a quantitative expression of the double standard in sexual morality. *Caring for* also shows a difference of 22 points. In this case the men give the higher score. Next in size of the difference comes *despise* with 18 points, followed by *pride* with 17. In both these cases also the men

[1] It is noteworthy that one of the strongest of the stimulus words, *supplicate*, owes its high position not entirely to the emotional value in the sense implied above, but rather to the fact that its meaning was not known to 38 of the subjects! These were divided equally among males and females. This entirely irrelevant circumstance gave rise to many "indicators" such as failure to respond, long reaction-times, clang association and failure to reproduce original reaction at second presentation. The value of *supplicate* as a word for touching emotional complexes is therefore questionable. Fortunately this is the only word of the 100 the meaning of which was not known.

yield the higher score. *Choosing* comes next, showing a difference of 16 points with the women once more in the lead. The reasons for the differences of the last four words while tempting to speculation, are by no means so easy to find as in the case of the word *fall*.

In a number of other cases where the writers expected to find marked sex differences, none large enough to be significant were found.

III.

The Relative Frequency of the Various Alleged Complex Indicators

The utility of a given type of complex sign depends chiefly upon two facts: (1) The relative frequency of the appearance of this sign and (2) the diagnostic power or reliability of the sign when it does appear. Clearly the sign of an unfailing indicator is of little value if it is of rare occurrence and a sign no matter of how frequent occurrence, is of no value if it is nearly as likely to fail to indicate as to indicate truly.

The number of times each kind of indicator appeared with each of the 100 subjects was therefore tabulated by sexes, totaled and averaged. The final distribution of the 6,639 indicators yielded by the experiment is shown in Table II. The subjects average slightly over 66 complex signs each. Long reaction-time is the most numerous of all the signs, yielding 20.3 per subject. This is followed closely by repeated use of the same response word, and quite closely by defective reproduction of the response word. Next in importance but far less frequent comes repetition of the stimulus word with about one per subject. Failures to respond may probably be classed as unduly long reaction-times while the two remaining types are too infrequent of occurrence with the present group of subjects to yield reliable results from an analysis.

The sexes are remarkably equal in the total number of indicators, there being but one point difference between the totals of the two groups. But here the similarity largely ceases, as appears from an examination of the first two columns of Table II. The women show 92 more cases of long reaction-time than the men, which is about 10 per cent. excess. In the case of failure to respond, which probably may be regarded

TABLE II.

Type of Indicator	Total Number of Appearances in:			Average Number of Appearances per Subject Irrespective of Sex
	50 Men	50 Women	Both Men and Women	
Long reaction-time	973	1,065	2,038	20.38
No response	20	30	50	.50
Repetition of stimulus word	187	266	453	4.53
Assimilation of stimulus word	42	60	102	1.02
Defective reproduction	1,094	820	1,914	19.14
Repeated use of the same response word	964	1,051	2,015	20.15
Strange or senseless reactions	34	26	60	.60
Obvious perseveration of idea	5	2	7	.07
Total	3,319	3,320	6,639	66.39

as an extreme form of long reaction-time, the women show about 50 per cent. excess. This agrees with the finding of Wells[1] that the women are prone to give long reaction-times, though the present writers are uncertain about attributing it to the fact that the experimenter was a man, as does Wells. If this were the case ought we not to expect the women for the same reason to show an excess in the other signs as well? As seen above this is not the case. Again the women show an excess of almost equal proportions in the repeated use of the same response word. Paradoxically enough these two types of indicators have a distinct *negative* correlation with each other (Table IV.). This means that repeated use of the same response word is associated with *short* reaction-times, rather than long.

But the greatest difference of all lies in the defective reproduction of the response word, where the men show 274 more signs than the women. This is an average difference of about five per person and amounts to over 25 per cent. excess. It is interesting to speculate as to why the men show this indicator so much more frequently than the women. A

[1] Wells, F. L., 'Some Properties of the Free Association Time,' *Psychol. Rev.*, 1911, Vol. XVIII., p. 7.

possible explanation may lie in the commonly observed fact that women, on the whole, tend to be more careful and conscientious in their reactions to all kinds of experiments, and especially show a greater interest in making a good personal score. It is not hard to see how the greater carelessness of the male subjects might lead to a greater number of false reproductions. At the same time the greater preoccupation of the female subjects with making what they feel to be appropiate responses may quite probably explain the excess of long reaction-times previously noted.

One of the largest proportional differences is in the case of repetition of the stimulus word. Here the women exceed the men by about 35 per cent. As pointed out by Jones,[1] this doubtless is a form of 'stalling' or defense reaction. The present writers incline to regard it as a peculiarly characteristic difference in the behavior of the two sexes. Because repeating the stimulus word consumes time, this fact would also account for nearly all the excess of long reactions among the female subjects. This excess of long reaction-time was 92, while the excess of repetition of the stimulus word was 79.

IV.

The Diagnostic Reliability of the Various Alleged Complex Signs

It has been previously pointed out that the value of a particular type of complex indicator depends not only on the frequency of its appearance but also upon its diagnostic reliability when it does appear. That the potency of the different indicators should vary from one another is almost self evident. Despite the apparent clinical value of such knowledge, neither Jung nor any other investigator so far as the present writers have been able to discover, has seriously raised the question as to the relative diagnostic efficiency of the various signs. From the general clinical accounts of the method[2] it would appear that all of the indicators are given about equal weight, with the exception that special impor-

[1] Jones, *op. cit.*, p. 407.

[2] Jones, *op. cit.*

tance is sometimes attached to lengthened reaction-time. In the present section we shall therefore attempt (1) to get some indication as to which of the various alleged complex signs are in truth associated with emotional complexes and (2) to secure some evidence as to the relative diagnostic power of those signs which appear to be valid.

The method of attack is almost of necessity an indirect one because of the presumably subconscious nature of some of the complexes and of the impossibility of obtaining honest introspections in many cases on other complexes which may be more or less clearly conscious. Stated in axiomatic form, the principle underlying the method is that, other things being equal, two variables related positively to the same variable are positively related to each other. For example, if all reaction-times of over thirteen fifths of a second were associated with emotional complexes, and all times shorter than this never were; and if in addition, defective reproductions always were associated with complexes, while successful reproductions never were; then of necessity, long reaction-times must always accompany defective reproductions and short reaction-times, successful reproductions. That is to say, there would be a perfect association between long reaction-times and defective reproduction because of their mutual relation to the complexes, and entirely apart from any natural affinity between the two indicators arising from their own proper natures. In such a case the degree or closeness of their association with each other would serve as a measure of their closeness of association with the complexes, *i.e.*, of their diagnostic reliability. If, however, upon computation we find that the degree of association between two alleged complex indicators is zero, we must conclude by similar reasoning and other things being equal, that one of them at least is not an indicator at all.

An excellent concrete illustration of this principle in a relatively clear form is found in the relation of long reaction-time to a number of other alleged indicators taken together. If long reaction-time on the one hand, and the several signs on the other, both indicate emotional complexes, then they ought both to tend to fall together upon such individual stimulus words as touch complexes. Accordingly words which yield long reaction-times should on the whole show a considerably larger number of other complex signs than words with short reaction-times. This is, in fact, found to be the case when the number of the other indicators per 100 stimulus words is represented graphically for the various lengths of reaction-time (see heavy curve, Fig. 3). Both the principle under consideration and the excellent reputation of long reaction-time as a complex indicator[1] thus find support and confirmation in the marked and consistent rise of this curve with increasing length of reaction-time.

A limitation of the above method of measuring the diagnostic potency of the various types of complex signs remains to be pointed out. It lies in the possibility that tendencies to positive or negative association among the individual indicators may rise from other sources than the alleged mutual association between the individual indicators and the complexes. However, with a group of several indicators, the various possible *individual* affinities and repulsions which may exist among them apart from a common positive relation to the complexes, ought largely to neutralize each other. Thus the strength of the association found ultimately to exist between a given sign and *all other signs taken together* may be presumed to yield a useful indication of the closeness of association of the alleged sign with emotional complexes, and hence of its diagnostic reliability.

Jung himself recognized this principle[2] and made a limited use of it. He gives, for example, a proof based upon 2717 reactions from a variety of pathological and normal subjects, that defective reproduction of the reaction word is in truth a complex sign.[3] His proof however, though numerical, is qualitative rather than quantitative in its nature. While he shows that defective reproductions tend to be associated positively with long reaction-times, it seems not to have occurred to him to determine the strength or diagnostic power of this tendency. Yet he recognized that the reliability of

[1] Jones, *op. cit.*, p. 406.

[2] Jung, *op. cit.*, p. 188 ff.

[3] Jung, *op. cit.*, p. 191.

this indicator is distinctly limited. Fortunately one of his tables used in the above demonstration is so constructed that we were able to make the necessary computation for comparison with our own results.[1]

The method of computing the extent of association used in the present study may be illustrated by applying it to the relation of long reaction-times and the group of other complex signs just considered. The curve, as has been seen (Fig. 3), reveals a distinct tendency to association, but the *strength* of the tendency is not apparent. This may readily be determined by the following formula:[2]

$$r = \cos \frac{\sqrt{bc}}{\sqrt{ad} + \sqrt{bc}} \pi,$$

where r is a function comparable with Pearson's product-

TABLE III

FOUR-FOLD TABLE SHOWING RELATION OF THE LENGTH OF REACTION-TIME TO ALL THE OTHER INDICATORS LISTED ON PAGE 4 WITH THE EXCEPTION OF REPEATED USE OF THE SAME REACTION WORD

	Words Having One or More Other Indicators Exclusive of Long Reaction-times	Words Having no Other Indicators
Words having a reaction-time of more than 13-fifths of a second	(*a*) 840	(*b*) 1,200
Words having a reaction-time of 13-fifths of of a second or less	(*c*) 1,510	(*d*) 6,450

[1] Jung expresses the principle as follows: "Sind nun diese Merkale wirklich bezeichnend d.h. hat die analytische Methode heir zu einem richigen und nachprüfbaren Resultat geführt, dann müssen die Merkmale im allgemeinen unter sich in nahen Beziehungen stehen, d.h. sich bei gewissen Assoziationen mit Prädilektion begegen also z.B. mangelhalfte Reproduktionen und zu lange Zeiten. Ist das nicht der Fall, und zerstruen sich die Kompiexmerkale wahllos über den ganzen Versuch, dann hat allerdings die Analyse zu einem Fehlschluss geführt." The computation from Jung's table mentioned above, shows that there was an association of +.37 between long reaction-times and imperfect reproduction. This compares fairly well with the +.26 from our own results, when we consider that his subjects were largely pathological.

[2] Whipple, 'Mental and Physical Tests, Simpler Processes,' p. 48. While probably not so perfect, theoretically, as the more elaborate method of Pearson (Davenport, C. B., 'Statistical Methods,' p. 49 ff.) its use was more than justified by the enormous saving in labor of computation in a study already extremely laborious in this particular.

moments coefficient of correlation, π is 180° and the values of a, b, c and d are the respective entries in the four-fold table shown on page 125 (Table III.).

Substituting the values of the table in the formula, we have:

$$
\begin{aligned}
r &= \cos \frac{\sqrt{1,510 \times 1,200}}{\sqrt{840 \times 6,450}\ \sqrt{1,510 \times 1,200}}\,\pi \\
&= \cos .366 \times 180^\circ \\
&= +\,.408.
\end{aligned}
$$

In round numbers then, + .41 is the strength of the tendency to association between long reaction-times and the presence of the other alleged complex signs, shown graphically in Fig. 3. While by no means perfect (1.00) it is definite and significant.

The probable error computed by the rough formula:

$$
\begin{aligned}
\text{P. E.}_{.r} &= \frac{1.1}{\sqrt{n}} \\
&= \frac{1.1}{\sqrt{10,000}} \\
&= .011
\end{aligned}
$$

is probably not much in excess of .02. Since the n of all the other coefficients of association computed in this section is the same, .02 may be considered approximately the probable error of them all.

Proceeding as above, the coefficients of association were found between each of the five types of indicators which occurred frequently enough to give results (Table II.) and all of the other indicators massed together without distinction. These results are shown in column VI. of Table IV. It is seen that the repetition of the stimulus word is the strongest of all, giving a coefficient of + .53. Assimilation of stimulus word comes next with +.33. Long reaction-time comes third with + .24, defective reproduction is fourth with +.17, while repeated use of the same stimulus word comes last with the slight negative relation of − .06.

If taken at face value certain tentative conclusions may be drawn from these figures.

1. Repetition of the stimulus word is decidedly the most reliable diagnostic sign of the five indicators examined.

2. The first four are in all probability real complex signs.

3. Repeated use of the same reaction word is a complex sign of very doubtful diagnostic value, at least in the sense that the other signs are diagnostic. Ordinarily the appearance of the complex sign is supposed to apply to the stimulus word, meaning that the stimulus word has touched an emotional complex. The present evidence seems to indicate that repeated use of the same reaction word is somewhat more likely *not* to do this than it is to do so. There remains of course the possibility that the reaction may itself be symptomatic of a complex.[1] The present evidence has little or no bearing upon that question.

Table IV.

Showing the Coefficients of Association between the Various Signs and Groups of Signs.

	Repetition of the Stimulus Word	Assimilation of Stimulus Word	Long Reaction-time	Defective Reproduction of Reaction Word	Repeated Use of Same Reaction Word	All Other Indicators Massed Together	All Other Indicators Except Repeated Use of Same Reaction Word
	I	II	III	IV	V	VI	VII
Repetition of the stimulus word		+.41	+.69	+.14	+.24	+.53	+.586
Assimilation of the stimulus word	+.41		+.52	+.06	−.31	+.326	+.469
Long reaction-time	+.69	+.52		+.26	−.087	+.242	+.408
Defective reproduction of reaction word	+.14	+.06	+.26		+.017	+.174	+.259
Repeated use of the same reaction word	+.24	−.31	−.087	+.017		−.06	−.06

[1] Jones, *op. cit.*, p. 407.

It is a little surprising that long reaction-time should not show up any stronger in the series than it does. Indeed it would not have been beyond reasonable expectation to find it ranking first, in view of its great reputation as a complex indicator. It actually ranks third with a coefficient of + .24, which is less than half as strong as that of the repetition of the stimulus word. The present evidence, while not final owing to the limitations of the method pointed out above (p. 124), at least raises some presumption against the preeminence among complex signs now enjoyed by long reaction-time.

In order to analyze in further detail the origin and inner constitution of the above tendencies to association, the coefficients of association were computed between each of the above five indicators among themselves, by pairs, in all the possible combinations. These appear in the first five columns of Table IV. Here we may expect the individual affinities and repulsions among the various indicators to appear with some clearness. The strongest of these affinities (with a coefficient of + .69) between long reaction-time and repetition of the stimulus word, is obviously due in part to the fact that the repetition of the stimulus word itself takes time, and directly contributes to increasing the length of the reaction. The next highest is between assimilation of the stimulus word and long reaction-time, with a + .52. Perhaps this may be attributed to a certain amount of real misunderstanding of the stimulus word with the consequent accompaniment of a period of doubt before the decision (and so the reaction) could be made. A similar mechanism tends to account for the correlation of + .41 between assimilation and repetition of the stimulus word, since in the real uncertainty of understanding a word, we naturally tend to repeat it in a questioning tone to verify our interpretation.

In all of the cases of negative or practically zero correlation, we find involved the repeated use of the same reaction word. This tends to confirm our doubts expressed above (p. 127) of the diagnostic reliability of this alleged sign. The − .087 shows that repeated use of the same reaction word is, upon the whole, positively associated with *short* times rather than with long, though it might conceivably be associated with both (as contrasted with medium time) provided the line of regression were curved in a certain manner. This possibly will be examined in detail below. The only positive coefficient of any size involving repeated use of the same reaction word is with repetition of the stimulus word, + .24. This is extremely interesting, but not easily explained by the available facts.

In view of the very doubtful showing made by repeated use of the same reaction word, it was thought that avoidable distortions might have entered into the results in column VI. from its inclusion in the massed indicators. Therefore correlations were computed anew between each of the four apparently significant signs and all of the other indicators massed together except repeated use of the same reaction word. These results are given in column VII. It is seen at once that in all cases there is a marked increase in the degree of association, though the relative strength of the various indicators remains the same.

V.

Diagnostic Potency as a Function of the Length of Reaction-time

Instead of being of the presence-absence type of variable as are most of the other alleged complex indicators examined above, reaction-time is a continuous variable and for that reason offers certain advantages in the way of further analysis. Thus far we have tacitly assumed that all reaction-times longer than thirteen-fifths of a second indicate emotional complexes, while all shorter times do not. We now know (p. 127) that this is not always true, though a correlation of + .24 indicates that it is more apt to be true than false. The question still remains: At what point on increasing reaction-times do they first begin, upon the whole, to indicate emotional complexes? Is there a sharp cleavage at this point between the critical and the non-critical times, or is the transition gradual and continuous?

And what about short reaction-times? It has been stated by several writers that very short reaction-times also indicate emotional complexes.[1] Watson, in discussing this matter, says,[2] "The indicators of implicit response or tension obtained from the subject are unduly long reactions . . .; significant response words . . .; *too rapid responses* . . .; low level responses; failure in response . . ." (italics ours). To be sure the decided positive correlation found above between long reaction-time and the other indicators massed together, means superficially that there is *negative* correlation between those indicators and short times. The question can not however, be answered in this summary manner, for correlation coefficients are based upon the assumptiom of a straight line of regression. In the present case, on the contrary, it is conceivable that the line of regression might be in the shape of a very broad and low U with one upright arm somewhat higher than the other.[3] In such a case both extremes of time would be complex indicators, as contrasted with the intermediate times, though the long times would be stronger indicators than the short.

The obvious way to settle such questions as those raised above is to plot the actual curve of the relation in detail. This consequently was done. The indicators of the 10,000 reactions were tabulated according to the absolute time of each. We then found the average number of complex signs per hundred words falling on each length of absolute time from the longest to the shortest. These averages were based on groups varying from 500 to 1,000 reactions. They were then plotted and appear as the continuous line, Fig. 2. In order to locate the point at which increasing reaction-times become critical, a horizontal line was drawn at such a level that it would represent the number of indicators per hundred words if the indicators were distributed indifferently throughout all lengths of time. Wherever the curve rises above this line, then, at that point it begins to exceed a chance frequency,

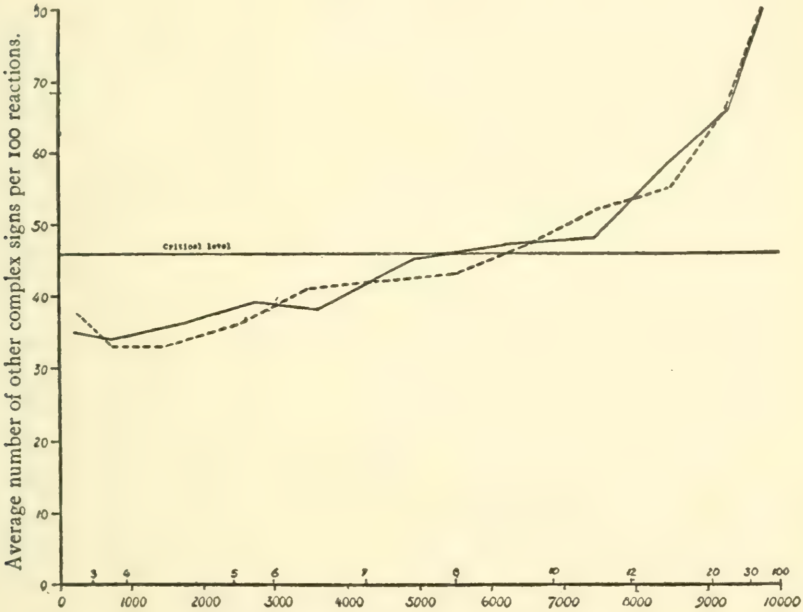
[1] Dooley, 'Lucile,' *op. cit.*, p. 132.

[2] Watson, J. B., 'Psychology from the Standpoint of a Behaviorist,' p. 208.

[3] Yule, G. U., 'An Introduction to the Theory of Statistics,' p. 162f.

i.e., it becomes significant; and the higher it rises the more significant it becomes.

An inspection of this curve shows that the critical point on increasing reaction-time, falls not far from nine fifths of a second. The excess above chance does not become sufficient to be practically significant, however, until about thirteen or fourteen fifths of a second, which agrees very well with current



Order of reactions from shortest to longest.

——————— According to the absolute method.

— — — — According to the relative method.

FIG. 2. Curves showing the relation of all the other complex signs massed together, to length of reaction-time. The figures *above* the base line represent length of reaction-time in fifth-seconds.

practice.[1] This curve also shows very clearly that there is no sharp division between the critical and non-critical reaction-times but that they pass over from one to the other by a most gradual and continuous movement. The only suggestion of such a cleavage lies in the fairly marked upward bend in the curve at about twelve fifths of a second which

[1] Brill, *op. cit.*, p. 140.

continues at a positive acceleration until the end. When plotted with a base line strictly according to absolute units of the time however, and not according to the temporal ranking of the reactions themselves as in Fig. 2, the curve becomes one of negative acceleration, *i.e.*, tending to the horizontal.

Lastly, the curve gives scant support for the claims made on behalf of short reaction-time. It is true that there is a suggestion of a secondary rise at the lower tip of the curve, and while probably a real tendency as we shall see later, it does not approach even remotely the level of diagnostic significance.

Since the above curve was based on the absolute reaction-time and since subjects differ greatly in the general rate of their reactions, many fast subjects' records are not represented at all on the right of the above curve while many slow subjects' records are not represented on the left end of the curve. The question naturally arises as to whether a method based on absolute times is as diagnostically significant as one based upon *relative* times. Clearly if an analogus curve were plotted according to the relative method and the downward slope from right to left should be found more steep in the latter case, the relative method would be shown the more significant. Accordingly each of the 100 reactions of each of the 100 subjects were tabulated anew, this time relatively. The number of indicators per 100 reactions was found for the five longest reactions of each subject (500 reactions in all), for the next five longest reactions, for the next ten longest (1,000 in all), the next ten and so on until at the short reaction-end, we descended the last two steps by 5's again. The two short steps of 5's at each end of the series, were made with the purpose of securing a little more detail at those especially interesting points. The curve appears as the broken line, Fig. 2.

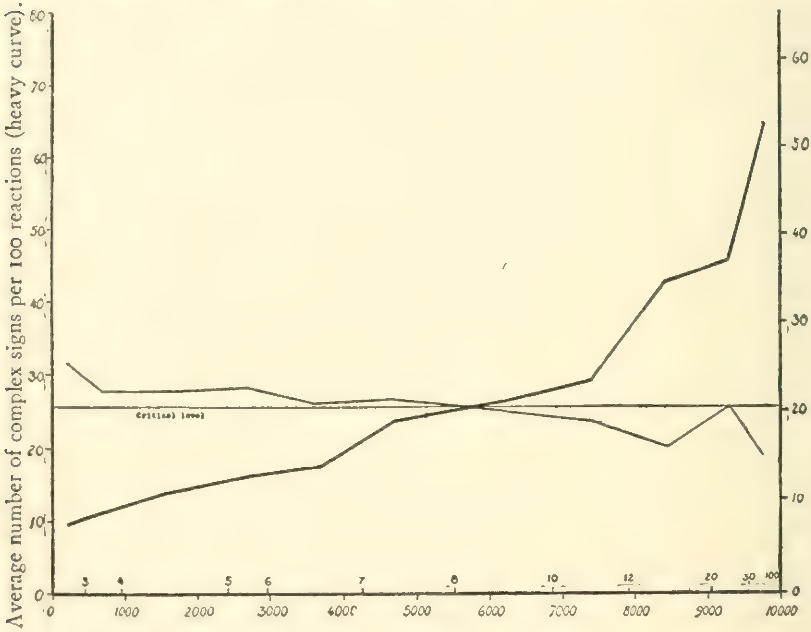
When we consider the marked difference in the actual data represented in the corresponding parts of the two curves and especially at the ends, they appear astonishingly similar. This detailed similarity can only mean that on the long run at least, one method is diagnostically as potent as the other. It is evident, then, that in so far as two methods differ, any potency peculiar to one method must be duplicated by an equal potency peculiar to the other,[1] and that the two individual potencies must follow the same law. No doubt this law is essentially the same as that of the central potency shared by both methods, *i.e.*, waxing in its progress from left to right in Fig. 2. Accordingly a combination method which would give joint weight to each ought to yield a higher diagnostic value than either method alone. This depends in the last analysis upon the fact shown by the above curves, that where one method includes more reactions from a given subject than the other, there exists on the long run, a significant potency in the longest of these reactions. For example, the above curve shows that the limit of practical diagnostic significance is about the fifteenth from the longest reaction. By the absolute method it is not far from thirteen fifths of a second. If in a given subject there were 35 reactions longer than thirteen fifths of a second and by the relative method only 15 reactions could be included, the joint method would make an equal concession to each method and take the longest 25. In another case there might be but one reaction longer than thirteen fifths of a second yet the relative method would demand 15. The joint method would then use the longest 8 reactions. It would be a relatively simple though somewhat laborious task to test empirically the truth of this theory.[2]

It will be noted that at the lower end of the broken curve we find the same upward rise previously noted in the other, only in this case it is much more marked. In order to get further light on this matter, the data was plotted for the men and the women separately and the rise was found in both alike. Without doubt it is a real tendency. Later we shall find an explanation for it.

[1] Partial correlations (Yule, *op. cit.*, pp. 239 ff) show that only about one third of the potency of each is shared by the other.

[2] This may be done rather easily, however, by Yule's method of multiple correlation (Yule, *op. cit.*, p. 248). This indicates that by combining the two methods as indicated above, the potency will be increased about 20 per cent. over either method independently.

The negative association of .087 obtained above between repeated use of the same response word and long reaction-times, raises the question as to the exact relation of these two alleged signs. As has been suggested in another case, it might well be that here also a curved line of regression would be found such that repeated use of the same reaction word would correlate both with short and to a lesser extent but positively also, with long reaction-time. To determine this the



Orders of reactions arranged from shortest to longest.

——— Repeated use of the same response word, absolute method.

▬▬▬ All of the other alleged complex signs except repeated use of the same response word, absolute method.

FIG. 3. Showing the relation of various complex signs to length of reaction-time. The figures *above* the base line represent length of reaction-time in fifth-seconds.

entire 10,000 reactions were again tabulated according to time to determine the exact relation of repeated use of the same reaction word to long and short reaction-time. The curve was plotted according to the same plan as that in Fig. 2. It is the lighter curve, Fig. 3. It shows at once a distinct and consistent correlation with short times. For the most part, it is surprisingly smooth and straight.

By subtracting the various parts of the data for this curve from the corresponding parts of the data for the original curve of Fig. 2, we obtain data for the plotting of the curve of the relation of the four other indicators, to reaction-time. It will be recalled that this relation yielded an association coefficient of + .408. The curve appears as the heavy line in Fig. 3. Here we find the curve distinctly steeper than the corresponding curve of Fig. 2. The rise at the lower end of the curve in Fig. 2 has also entirely disappeared. There is evidently no tendency whatever for these indicators to be connected with short reaction-time. The rise at the lower end of the original curve was evidently due entirely to the influence of repeated use.

Do the above results mean that short reaction-times and the repeated use of the same response word are both to be excluded from the list of complex indicators? So far as the present method can decide the matter with normal subjects, the writers are inclined to answer in the affirmative. It is true, these two alleged indicators are mutually related to the extent of + .087. While this coefficient is doubtless real in the sense that it is several times that of the probable error, it still is very small and probably no more than a natural individual affinity between the two types of reaction. There remains of course the possibility that the two alleged signs may be a weakly allied diagnostic group, significant with respect to the reaction word itself and not to the stimulus word, as is assumed to be the case with the other signs. The answer to this question must await further investigation.

VI.

The Diagnostic Potency of One Indicator vs. Two

In psychanalytic literature the notion is sometimes expressed or implied that when a stimulus word evokes a single complex sign, no particular significance is to be attached to it; but if two or more signs appear, the combination at once becomes decidedly significant.[1] This necessarily means that

[1] Dooley, *op. cit.*, p. 148.

the second indicator makes a very much greater increment to the diagnostic potency than the first. The question therefore arises as to how much if any increase in the diagnostic significance is added to one indicator by the presence of a second?

The indirect method utilized so largely above furnishes us with a ready means of getting at least an approximate answer to this question. To illustrate, a tabulation of certain data for this purpose showed that long reaction-time and defective reproduction falling together on the same stimulus word, attract the repetition of the stimulus word, 18.1 per cent. of the cases. On the other hand long reaction-time separately, attracted it 16.6 per cent. of the times, while defective reproduction attracted it 7.3 per cent. of the times, the two independently making an average of 12.7 per cent. Thus the two together show a gain of 5.4 points or 42½ per cent. over the two separately. From the data of the four strongest indicators investigated above, it is possible to find twelve different combinations such as just described. These differ much, as might be expected, but a weighted average of them all, probably does not depart very far from the truth. The computation shows that two given indicators are 64.8 per cent. more likely to attract a given third indicator than either one of them is likely to attract it separately. This suggests that while two indicators are distinctly more significant than one, the second indicator adds by no means as much diagnostic potency as the first, to say nothing of more.

The increment in diagnostic potency contributed by the addition of a third or fourth indicator, while of great interest and some practical importance, was not investigated. Presumably the increase in diagnostic potency resulting from the addition of successive indicators would follow a law of diminishing returns.

EFFECTS SIMULATING FATIGUE IN SIMPLE REACTIONS

From the Psychological Laboratory of McLean Hospital[1]

F. L. WELLS, Ph.D., In Charge
C. M. KELLEY, M.D., Associate
GARDNER MURPHY, A.M., Assistant

1. The time of the simple reaction process has generally been regarded as very resistant to fatigue. Observations by Cattell have given the chief ground for this view, but his classical experiments (1) on this subject hardly show such absence of fatigue loss as to justify regarding the question as closed. There were studied discriminative reactions to light, letters and other associations, as well as simple reactions to light and sound. Observations covering periods of about 16 hours were made with two subjects. Signs of fatigue are variable, but certainly present in various series. The final light reaction series for Subject *C* is 122 per cent. of the initial reactions series; in Subject *B* this fatigue effect is hardly traceable. Both subjects show moderately increased times for letters, and for the associations the final time averages in *B* 131 per cent. of the initial, for *C* 113 per cent. Simple reaction time to sound cannot be said to show a loss in either subject. It is observable that variability of reaction times does not change in proportion to the length of the reaction times themselves; also that there is generally more loss on the visual side, where the chances of sensory fatigue are greater; also that there is more fatigue loss in the more automatic functions, where it is less offset by practise.

2. In the series reported by Bettmann (2), of 1,000 (apparently to sound) choice reactions made in two hours, the score for the final fifth is 108 per cent. of the initial fifth. Woodrow (3) reports shorter reaction times obtained at the end of an hour's reaction time work of similar character to the present,

[1] As of August 1, 1920.

but under conditions more generally stimulating to the subject.

The present experimental facts are soon told. When 216 reactions are taken under the conditions of these experiments (4), the reaction time in the second 108 reactions averages close to 105 per cent. of that in the first 108, as appears in the following figures:

PER CENT. WHICH THE REACTION TIME SCORE IN THE SECOND HALF OF THE EXPERIMENT IS OF THE SCORE IN THE FIRST HALF

Subjects	A-group (13 Persons)	*K*	*W*
Av.	104.9	105.1	106.3
M.v.	4.9	4.8	3.5

3. Variability within each unit reaction-series (27 reactions) is indicated by the per cent. which its *upper* quartile is of its median. This figure is termed the 'quartile ratio.' It is observed that while the central tendencies are longer in the second half (par. 2), this lengthening is unaccompanied by corresponding increase in the quartile ratios. These are quite stable, as follows:

PER CENT. WHICH THE QUARTILE RATIOS OF THE FIRST HALF ARE OF THE QUARTILE RATIOS OF THE ENTIRE EXPERIMENT

Subjects	A-group	*K*	*W*
Av.	49.9	50.2	49.8
M.v.	1.2	.7	.4

4. It runs counter to the weight of evidence to suppose the loss due to actual effort in the repeated reactions themselves. The series last some forty minutes each, and there is a recovery period between single observations of about eleven seconds. Appeal can hardly be taken to exhausting character of the work itself.

5. More suitable interpretation is found in certain special conditions of the work. The room in which it takes place is considerably darkened, to permit reading of a galvanometer scale. The subject's only sense stimuli incidental to the experiment are the closing of the experimenter's key for the get-ready signal, the get-ready signal itself, the stimulus for reaction, and whatever noise or other disturbance the subject

makes in his own responses. The galvanometer chronoscope operates entirely without noise. The average interval between stimuli is slightly over 11 seconds, which makes the reactions seem very discontinuous. As an experiment progresses, the subject is apt to show signs of increasing somnolence, yawning or stretching, etc. The writer (*W*), found it the most *ennuyant* of experiments, as the doctor of physick his own draught, until it became possible to perform it 'from the spinal cord.' Part of this 'logeyness' induced by the conditions, seems to be a slightly increased refractoriness to the simple reaction process. It is well to note that while the reactions become definitely longer, they do not become more variable (par. 3). The loss does not involve a diminution of attentional control, as expressed in tendency of attention to occasional lapse. It is rather an upward thrust of the whole distribution.

6. There appears in these experiments a decrease in functional efficiency under continued exercise of the function. The question if fatigue is a term properly applicable to such decrease leads into a somewhat casuistical discussion. Thorndike (5) has regarded it as a characteristic of fatigue, that it shall be a condition improvable by rest. It is improbable that the phenomena meet this criterion, as rest is ordinarily conceived. A lessening or temporary cessation of the experimental work, or a further lessening of the already slight general stimulation which the environment of the subject affords, would be likely to further decrease his efficiency, or abolish it entirely by putting him to sleep. The 'rest' that would cure this decreased efficiency would be some relief from the rather depressing experimental conditions described.

7. The writers favor regarding this slowing of reaction time as a fatigue effect, under the concept of fatigue as a name including all influences by which the continuance of a task tends to decrease its efficiency. Accumulation of fatigue toxins as result of either mental or physical overexertion seems to be here ruled out. Secondary effects of disagreeable sensations from such exertions seem also ruled out. The effective

factor is rather a 'monotony,' and its result, a raising of the reaction threshold.

8. It may be asked if there is any difference of sound and light stimuli in this respect, as these sensations differ so much in their own fatigue phenomena. A generally negative answer is implicit in the findings of a previous paper. In both halves of the experiment, the sound-light ratios (per cent. which the light time is of the sound time) are similar for the A-group of thirteen subjects and for Subject *W*. In Subject *K* the sound-light ratios are clearly less in the second half; that is, the sound reaction has lost *more* than the light. This is the opposite of what would be expected from a sensory fatigue effect.

9. A similar question may be asked regarding the two different prestimulus (get-ready) intervals employed; whether the monotony and consequent 'logeyness' have the more effect on a one-second or a three-second prestimulus interval. As expressed in the A-group, the general effect is that the monotony is especially unfavorable to the longer, three-second, prestimulus interval. The average difference in favor of the one-second interval in the second half, is the same for sound and light. The differential effect on the prestimulus intervals does not appear in *K* and *W*, there being substantial equality throughout their records. This should be taken in connection with the presumably lesser conscious attention in these subjects during their repeated series. In *W* especially, the process was little short of automatic, with no conscious attention to the interval employed.

For the A-group, the longer times in the second half are essentially due to the results with the three-second interval. The difference for the one-second interval is hardly significant. The figures are as follows, in which there again appears the tendency to smaller variations in the results for light:

Per Cent. which the Scores for the Two Prestimulus Intervals in the Second Half are of Their Scores in the First Half; (A-Group Only)

	Sound, 1″	Sound, 3″	Light, 1″	Light, 3″
Av.	102.7	108.1	101.2	107.0
M.v.	7.3	11.5	5.9	3.9

10. In respect to the general disfavoring in the A-group of the longer interval by the monotony, the indication is that the 'fatigue' of monotony makes it little if at all harder to *rouse* the attention (to a one-second interval), but rather harder to *hold* it (to a three-second interval). This is the early monotony effect. With the subjects accustomed to experimental conditions, *K* and *W*, the monotony effect is more evenly distributed to both prestimulus intervals.

11. It is a fate of concepts either to die barren, or to perish in bringing forth new ideas. 'Secondary dementia,' the Freudian sexuality, the Binet intelligence, herein bear the better part. A similar destiny perhaps awaits the concept of fatigue. The fundamental implications of the term are two: continued subjection to certain working conditions, and decreased effectiveness of *some* response. There are various formulations of how the first of these brings about the second. They are given in terms of toxins, reflex inhibitions, refractory phase. Monotony has also been cited as a factor in fatigue effects. The word implies a feeling of disagreeableness resulting from sameness. It is reasonable to regard this disagreeableness as secondary to an increasing *difficulty resulting from sameness*, as well as from general lack of environmental stimulation, and to include it in a broad conception of 'refractory phase.' It may be noted that saying the reaction time is increased in the second half of the experiment is the same as saying that it requires a greater stimulus in the second half to keep the reaction time at its former level; the reaction threshold is by so much raised.

12. In sum, the present experimental conditions are such as to make the reaction times in the second half of a forty-minute experiment, average some 105 per cent. of those in the first half. These conditions, with their moderate work and ample rests pretty well exclude exhaustion in the lower nervous arcs and allow a monotony effect to appear in relatively pure form. A distinctive feature of it is selective action on the longer, three-second prestimulus interval. Again, it does not alter the variability of the reaction times, which showed some tendency to grow greater in the results

of Cattell. The alertness aroused by the get-ready signal is not so well maintained under monotony as at the beginning, but it does not fluctuate any more under monotony.

REFERENCES

1. Cattell, J. M. Psychometrische Untersuchungen. *Philos. Stud.*, 1886, **3**, 489–492.
2. Bettman, S. Beeinflussung einfacher psychischer Vorgänge durch körperliche und geistige Arbeit. *Psychol. Arb.*, 1896, **1**, p. 157.
3. Woodrow, H. The Effect upon Reaction Time of Variation in the Preparatory Interval. *Psychol. Monog.*, 1914, **17**, pp. 55, 57.
4. Wells, F. L., Kelley, C. M., Murphy, G. Comparative Simple Reactions to Light and Sound. This Journal.
5. Thorndike, E. L. The Curve of Work. *Psychol. Rev.*, 1912, **19**, 165–194.

'SUBJECTIVE' PERCEPTIONS

BY KARL M. DALLENBACH

Cornell University

TABLE OF CONTENTS

In an experiment published some time ago by C. A. Britz,[1] in which he asked his observers to describe a field of colored rectangles presented tachistoscopically, figures were occasionally reported which were not present in the physical stimulus. These figures were regarded by Britz as 'subjective,' and as such received no further consideration in his published paper.[2] Subsequently we repeated Britz's experiment[3] and our observers also gave 'subjective reports.'[4] In the exposition and report of this work we too excluded them from our classification[5] and considered and listed them separately.[6]

[1] C. A. Britz, 'Eine theoretische und experimentelle Untersuchung über den psychologischen Begriff der Klarheit,' Saarlouis, 1913.

[2] *Op. cit.*, 56.

[3] Karl M. Dallenbach, 'Attributive *vs.* Cognitive Clearness,' JOUR. OF EXP. PSYCHOL., III, 183–230.

[4] The terms 'objective' and 'subjective' are used without prejudice. 'Objective' is used to designate the cases in which there is a physical correlate on the stimulus card; and 'subjective' is used to designate the cases in which there is no known physical correlate.

[5] *Op. cit.*, 189.

[6] *Op. cit.*, 188–189, and Table II., 195.

Not, however, from a conviction that this procedure was justified, but wholly because in this part of the work, that is, in the analysis of the perceptual phases, we desired to reduplicate Britz's conditions and to follow as closely as possible his schema of classification. Not to have done so would have exposed us to the criticism that we had changed the criteria of classification; and, since we were able to extend his classification from 10 to 17 steps, might also have prompted the suspicion that the extension was due in all probability to this change.

Britz's justification for the elimination of the 'subjective reports' may, however, be seriously questioned. For, in the first place, the present-day psychology of perception admits that the distinction between sensory and imaginal elements is purely logical and not psychological; and, in the second place, the comparison of a perception with an exposure field can be made only upon the assumption that the stimulus is physical and not physiological. Furthermore, where, as in this case, we are concerned with a problem of clearness, any question which might be asked concerning the sensory might also be raised concerning the imaginal elements. It would seem, therefore, that there is no *a priori* reason for the elimination of the 'subjective reports,' and if there is any *a posteriori* ground for such a procedure it can be found only in a study of the results themselves.

In the hope of answering these questions we have reexamined our original data, culled out all subjective reports, and subjected them to the same sort of analysis that the objective reports received in our earlier paper, and we have also compared at every stage the two sets of results.

Results

A. Cognitive Classification

In the cognitive classification of the subjective reports 10 of the 17 steps obtained in the earlier work were again obtained: 6 were found in *W*'s (Professor H. P. Weld's) reports; 7 in *F*'s (Professor W. S. Foster's); and 5 in *D*'s (the author's) reports.

The lowest step that we marked in the earlier work was one in which the observers described their impressions as 'spots,' 'flecks,' 'blurs,' without being able to say anything definite about form, size, quality or number. Only one of the observers, *W*, gave a subjective report of this nature. After an exposure of a card of Series II. (six rectangles) he accounted for six 'spots' and then continued, "All others—I guess there were one or more, I cannot say,—were vague."

The second step, in which a specific number of indefinite forms was reported, was identified in the subjective reports of all of the observers. Examples are:

W (7). (After an exposure of a Card of Series II., 6 rectangles.) "I saw seven areas."

F (50). (Card of Series II.) "My impression is of 7 areas."

D (78). (Card of Series II.) "Perception of 7 areas. Haven't the least idea of the quality of second. [The second was the subjective, for the descriptions of 1, 3, and the others were identical with the physical stimuli.] I cannot say whether the area was colored, or even whether it was rectangular."

The third step, in which a tint was ascribed to the indefinite impressions, did not appear in the subjective reports; but the fourth step, in which the impressions were described as colored, was found in the reports of two observers:

W (11). (Card of Series II.) "Blue, left to right . . .; red, to left of center . . .; five or six other colors were there."

D (19). (Card of Series II.) "Perception of 7 colored areas, B, P, O, —, —, —, —. [Blue, purple, and orange were the qualities of the three rectangles to the left of the card.] The areas to the right were unclear. Can report only that they were colored. I have no idea of the hue, tint, or chroma; did not perceive shape; know only that they were colored."

The fifth, sixth, seventh and eighth steps were not identified in the subjective reports of any of the observers.

The ninth step, in which a definite number of rectangular areas was cognized, was identified in the subjective reports of all observers. Examples are:

W (123). (Card of Series II., 6 rectangles.) "Blue . . ., yellow . . ., red . . . and four others."

(142). (Card of Series I., 5 rectangles.) "Green . . ., orange . . ., purple . . ., yellow . . ., and two others."

(163). (Card of Series II.) "Blue . . ., red . . ., yellow . . ., and three others."

F (15). (Card of Series I., 5 rectangles.) "Six rectangles, —, violet, —, gray, —, blue."

(40). (Card of Series II.) "My impression is of 8 rectangles, O, G, R, —, Bk, —, B, —; . . ."

(41). (Card of Series I.) "Yellow, red, green, —, —, blue. I think there were six colored rectangles."

D (116). (Card of Series II.) "Seven rectangular areas, named in order from left to right: Bk, B, R, Y, —, —, —. . . . Cannot report further about the quality of the three to the right; was only dimly aware of them."

(182). (Card of Series I.) "Six rectangles: B, G, B, Bk, —, —; . . . I am unable to add to my report of the last two."

(237). (Card of Series II.) "Seven rectangles: R, O, B, Y, —, —, —."

Isolated cases of the tenth step, in which a tint was ascribed to the rectangular forms, were obtained from all the observers:

W (97). (Card of Series II.) "Y . . ., P . . ., R . . ., and I should say four others. These appeared as grays, some darker than others."

F (10). (Card of Series I.) "Six rectangles: —, O, P, —, B, —. The last [the areas for O, P and B corresponded to the 2, 3 and 5 objective stimuli] was dark to medium tint."

D (64). (Card of Series II.) "Seven rectangles: R, Y, Gy, —, —, —, —. Have general impression that the unnamed areas were dark."

The bulk of the subjective reports was found at the eleventh step, in which the rectangular forms were described as colored: 62 per cent. of *W*'s and 78 per cent. of *F*'s reports were grouped here.

W (18). (Card of Series I.) "Six rectangles, blue last to right, good chroma, dark tint; red last to left, was also of good chroma and dark tint. There were four other colors which I cannot name."

(117). (Card of Series II.) "Seven rectangles: Bk . . ., B . . ., R . . ., and four other colors."

F (26). (Card of Series I.) "Impression of six colored rectangles. . . ."

(205). (Card of Series II.) "About 8 colored rectangles: Y, B, Bk, grey, purple, colored, colored, colored. . . ."

D (70). (Card of Series II.) "Eight rectangular areas, all colored. Cognized the colors from right to left: B, Bk, R, salmon. Haven't the least idea of the qualities of the other areas, although I know that they were colored."

(102). (Card of Series I.) "Perception of 6 colored rectangular areas. . . ."

The eleventh step was the highest of the cognitive phases that we were able to identify in the subjective reports of *W* and *D*. *F*, however, gave reports which could be classified at still higher levels. None was obtained from his reports at the twelfth, fifteenth and sixteenth steps, but several were obtained at the thirteenth step, and isolated cases were obtained at both the fourteenth and the seventeenth steps.

Examples of the cases at the thirteenth step, in which a specific quality was named, are:

F (32). (Card of Series I.) "Six rectangles: G, Bk, Y, P, B, Gy. I reported the colors to myself in order of reading and as recorded above; simply saw the colors clearly and at once said them, not, I think, on basis of memory after-image. Indeed when I came to ask myself the nature of any of the colors after reporting them I couldn't get a single color to pop up sensorial-fashion as they do in the memory after-image, or in 'immediate memory.'"

(117). (Card of Series II.) Named in his habitual way from center to left and then from center–right; "Grey, P, B, Bk, Y, —, grey. . . . The greys were of the same quality."

TABLE I

SHOWING THE VARIOUS COGNITIVE PHASES IDENTIFIED IN THE SUBJECTIVE REPORTS, THE NUMBER OF CASES OCCURRING AT EVERY PHASE, AND A DIVISION OF THESE CASES ACCORDING TO THEIR ATTRIBUTIVE CLEARNESS

Degrees of Cognition		Observer														
		W					*F*					*D*				
		Clearness				Total	Clearness				Total	Clearness				Total
		Not Reported	One Level	Two Levels			Not Reported	One Level	Two Levels			Not Reported	One Level	Two Levels		
			Upper	Upper	Lower			Upper	Upper	Lower			Upper	Upper	Lower	
Indefinite form																
Others experienced . .	1				1	1										
A definite number experienced	2				1	1				3	3	2			4	6
Tint	3															
Colored	4				14[1]	14									2	2
General quality	5															
Specific quality	6															
Hue, tint, chroma . . .	7															
Definite form																
Others experienced . .	8															
A definite number experienced	9	14	4		7	25				8	8	4	2		8	14
Tint	10		1			1				1	1		1		1	2
Colored	11	20	23		26	69	17	55		1	73		2		2	4
General quality	12															
Specific quality	13							6			6					
Hue and tint	14							1			1					
Hue and chroma	15															
Hue, tint and chroma	16															
Paper character	17							1			1					
Total		34	28		49	111	17	63		13	93	6	5	17		28

[1] One case included here was reported as at the third level; see earlier work, pp. 195, 225–227.

The report which we classified at the fourteenth level, in which a description of hue and tint was given, is:

F (98). (Card of Series II.) "Seven: B, Bk, BG, P, V, Y, dark grey." The objective stimuli were Bk, B, G, R, P, V. It is therefore obvious that the Y which *F* described further in his introspection as "light" was the subjective quality in this list.

The report placed at the seventeenth level, in which the paper character was described, is:

F (205). (Card of Series II.) "About 8 colors: Y, B, Bk, grey, P, —, —, —. . . . The Bk was not the velvet black, but the smooth one." The objective qualities were Y, B, grey, P, V, O; so it is again obvious that the Bk, which *F* described fully and which he inserted between B and grey, is subjective.

A résumé of the various cognitive phases which we were able to identify in the subjective reports is given in Table I., which shows in the 7, 12 and 17 columns, under the caption 'Total,' the total number of such reports given at the different cognitive levels by every observer.

A comparison of Table I. with the analogous table of the earlier work (Table II., p. 195) reveals, in spite of the fewness of the subjective cases involved, striking resemblances in the cognitive distribution of the two kinds of reports. These resemblances are more clearly shown in Tables II. and III. which follow.

Table II

Showing the Total Number of Times the Subjective Reports Were With and Without a Definite Form, and a Division According to Their Attributive Clearness

Observer	Report of Form	Attributive Clearness				Total
		No Report	One Level (Upper)	Two Levels		
				Upper	Lower	
W.	Indefinite.	0	0	0	16[1]	16
	Definite.	34	28	0	33	95
F.	Indefinite.	0	0	0	3	3
	Definite.	17	63	0	10	90
D.	Indefinite.	2	0	0	6	8
	Definite.	4	5	0	11	20

[1] One case included here was reported as at the third level; see earlier work, pp. 195, 225–227.

Table II. is analogous to Table III. (p. 196) of the earlier work. Both of these tables show that in comparatively few reports were the perceptions without definite form. The percentages of the cases which were with and which were without definite form for two of the observers, as is clearly shown in Table III., are nearly identical for the objective and the subjective reports.

TABLE III

SHOWING A COMPARISON OF THE 'FORM' OF THE PERCEPTION IN THE OBJECTIVE AND SUBJECTIVE REPORTS

Observer	Report of Form	Percentages	
		Objective Reports	Subjective Reports
W	Indefinite	14.6	14.4
	Definite	85.4	85.6
F	Indefinite	1.7	3.2
	Definite	98.3	96.8
D	Indefinite	5.3	28.6
	Definite	94.7	71.4

The reports of *W* and *F* show practically no difference in the 'form' division: 85.4 per cent. and 98.3 per cent. of the objective reports, and 85.6 per cent. and 96.8 per cent. of the subjective reports are of definite form for *W* and *F* respectively. *D's* results are not in such close agreement: 94.7 per cent. of the objective reports are 'definite' in form, whereas only 71.4 per cent. of the subjective reports are so distinguished. This discrepancy, however, is not large. The preponderance of cases in both kinds of reports is in the same direction—that is, in the direction of definiteness—and though the percentages are not identical, as they may be said to be in the cases of the two other observers, they are close enough, particularly when the small number of subjective cases reported by *D* is considered (which is probably due, not to the fact that he was a more objective type of observer, but to the fact that he had constructed the exposure-cards and was familiar with their schema and construction), to justify us in concluding that, from the point of view of definiteness and indefiniteness of form, the subjective and objective reports do not differ.

There is some indication, however, that the highest degrees of cognition are, except in rare and unusual cases, limited to the reports which have a definite and known objective correlate. *W* and *D* give no cases above the eleventh step, and *F* but two above the thirteenth. The apex of the curves of the objective reports (cf. Table I., p. 195, of the earlier work) lies for observers *W* and *D* at the sixteenth step, whereas the apex of the subjective reports (cf. Table I.) lies at the eleventh and ninth steps respectively for these observers. The shift in the subjective reports is five steps down for *W* and seven steps down for *D*. While this effect is not so pronounced in the results of *F*, the same general tendency is to be observed. His results in the 'objective' classification are bi-modular,—an apex occurs at the thirteenth and sixteenth steps. In the 'subjective' classification the apex is at the eleventh step, so that the same downward tendency is also evidenced in his data. Over 50 per cent. of the objective reports for *W* and *F* and over 75 per cent. of them for *D* are above the respective modes of the subjective reports. These results seem to indicate that the subjective reports are as a rule cognitively less clear than the objective.[1]

B. Attributive Classification

As we said in our earlier report, the observers were not always able, because of the complexity of the problem, to carry abreast the dual task laid upon them; that is, they were not always able to note meaning and process and to give a full account of both. They tended to concentrate either upon the fulfillment of the first part of their instructions,[2] in which case the description of process was vague and uncer-

[1] No one, in the work upon imagery and the differentiation of it from sensation, has ventured to suggest that the differences may in part be due to cognition. In view of the results of this study, though confessedly we were working at a higher psychological level than sensation—still in this connection it should be recalled that Britz thought he was working at the elementary level and chose his stimuli, the colored rectangles, with that idea in mind—we nevertheless hazard the opinion that cognition plays an important rôle in the differentiation of sensation and image.

[2] *Op. cit.*, 187.

tain, or on the introspection of process, in which case the statements of meaning were neglected and incomplete.[1]

Consequently, in some of the experiments, no mention was made of attributive clearness. These cases appear in the tables under the caption 'Clearness not Reported.' In other experiments clearness was reported at one level, at two levels, and by *W* in a few instances at three levels. There was, however, but one case in which a subjective report was given in a three-level consciousness. As a matter of economy of space we have grouped this case with those of the lower level of the dual formation; we are justified in this procedure by the fact that the three-level report proved upon analysis to be cognitive and not attributive. A note of the inclusion is made at the foot of all of the tables in which it appears.

1. *Clearness not Reported.*—Examples of the introspections in which subjective reports occurred, and in which the description of attributive clearness was omitted, are:

W (77). (Card of Series I., 5 rectangles.) "Orange, last to left, reddish, good chroma, medium tint; blue, next to orange, good chroma, medium tint; yellow, last to right, slightly greenish, poor chroma, very light tint; three other colors. I nearly forgot the O and B before I got them written down. The name 'orange' was kinæsthetically inhibited for some reason and for an instant all memory of the experience faded out. No clearness reports. I am certain of the descriptions however."

F (50). (Card of Series II., 6 rectangles.) "P, O, B, Y, and four other colors. My impression is of seven areas. All of the qualities named were particularized. B was dark but was not the Milton-Bradley blue; P was, I think, quite bluish; O and Y were ordinary, but were noted as being ordinary, not merely taken for granted or merely named. It seems very queer, but simply by the time I come to report I cannot say whether all the colors were on the upper level or not. I cannot even say whether there were differences of clearness among the four described."

D (217). (Card of Series II.) "Cognized from left to right; 1. red, good chroma, bright cherry red; 2. yellow, bright; 3. quality apprehended but cannot name it; 4. yellow, bright golden quality, brighter than that of 2; 5 and 6. dim awareness of these areas; 7. blue, cognized general hue, nothing else. Cognitively there were many degrees of clearness, but cannot report on the attributive clearness of these experiences."

2. *Clearness One Level.*—Examples of the introspections in which the subjective reports occurred, and in which the clearness was said to be of one level, the upper, are:

[1] *Op. cit.*, 196.

W (27). (Card of Series II.) "Yellow, cannot describe it further; blue, good blue, rather dark; green, good chroma, medium tint, I think; and four others. There must have been seven colors in all. . . . Total field was clear."

F (41). (Card of Series I.) "Y, R, G, —, —, B. I think there were 6 colored rectangles and that all were on upper level of clearness. . . ."

(44). (Again, since over two thirds of *F*'s reports, both objective and subjective, were of this nature.) "Seven colors, all at upper level."

D (221). (Card of Series II.) "Seven colored areas. Cognized from left to right: B, Bk, —, Gy, —, —, Y. B, merely the general hue; Bk, only a dark shade; Gy, medium, cognitively clearest of all, definitely bounded and a rough paper quality; Y, bright, without form, only a quality and tint. The sensations though equally clear were not maximally clear."

3. *Clearness Two Levels.*—Examples of the introspection in which subjective reports occurred, and in which the attributive clearness of the processes at the time of the exposure was said to be of two levels, are:

W (17). (Card of Series II.) "Blue, light tint, fair chroma; orange, medium tint, good chroma. Five other colors, seven in all. B and O were clearer during the exposure than the other colors."

F (152). (Card of Series II.) "Seven colored rectangles: B, —, —, G, Bk, P, —. G and Bk were clearer and held my attention longer than the others. Bk was a velvet Bk, G was rich, possibly very slightly yellowish, medium tint. P and B merely names and were at the same level as the others indicated by dashes."

D (231). (Card of Series II.) "Seven colored rectangles. From left to right they are Y, R, B, Bk, —, —, —. Those at the left clear; those at right unclear."

A summary of the attributive classification, just mentioned and illustrated above, is given in Table I. opposite the legend 'Total.' A comparison of this summary with the similar summary of Table II. of the earier study (*op. cit.*, p. 195) reveals a striking difference between the subjective and objective reports. This is shown in Table IV., in which the numerical summaries given in the tables mentioned above are presented for comparison in the common denominator of per cent.:

TABLE IV

SHOWING THE PERCENTAGES OF OBJECTIVE AND SUBJECTIVE REPORTS IN WHICH ATTRIBUTIVE CLEARNESS WAS NOT REPORTED; IN WHICH IT WAS REPORTED AT ONE LEVEL; AND IN WHICH IT WAS REPORTED AT EITHER THE UPPER OR THE LOWER LEVEL OF THE DUAL LEVEL FORMATION

Observer	Attributive Clearness	Percentage	
		Objective Reports	Subjective Reports
W	Not reported	34.9	30.6
	One level (Upper)	22.9	25.2
	Two levels: Upper	18.1[1]	....
	Lower	24.1	44.1[2]
F	Not reported	13.7	18.3
	One level (Upper)	73.7	67.7
	Two levels: Upper	8.0	....
	Lower	4.6	14.0
D	Not reported	9.8	21.4
	One level (Upper)	56.1	17.9
	Two levels: Upper	19.3	....
	Lower	14.8	60.7

The outstanding features of this table are: first, that no subjective reports are given from the upper level of clearness when the sensations aroused by the stimulus card are reported as of two levels; secondly, that there is a marked shift of the subjective reports towards the lower level. This is evinced in two ways: first by a large percentile increase, as shown by *W* and *D*, in the percentage of cases occurring at the lower level; and secondly by a large percentile increase, as shown by *F* and *D*, in the percentage of cases in which the clearness was not reported.

Of the cases in which clearness is given, 36.9 per cent. of the objective reports and 63.6 per cent. of the subjective are given by *W* from the lower clearness level; 5.3 per cent. of the objective and 17.1 per cent. of the subjective are given by *F* from the lower level; and 16.4 per cent. of the objective and

[1] *W* reported the attributive clearness of 5 experiments at 3 levels. The upper level was included here, the second and third levels were grouped with the "Lower."

[2] One case included here was reported as at the third, the lowest level of a three-level type of consciousness (*op. cit.*, p. 195).

77.2 per cent. of the subjective are given by *D* from the lower level. The percentage increase of reports at the lower level runs, for the subjective classification over that of the objective, from nearly 100 per cent. for *W* to over 300 per cent. and 400 per cent. for *F* and *D* respectively.

These data lead us to conclude that as a rule subjective are less likely than objective reports to be attributively clear.

Since the subjective reports are capable of the same classification as the objective, and since they do not show differences of kind but only differences of degree, we feel justified in concluding that Britz should not have omitted the subjective reports from his perceptual classification. Psychologically, these reports are just as good perceptions as those which have a known objective correlate.

C. Conditions of Occurrence

The second part of this paper is concerned with the questions where, when and under what conditions the subjective reports occur. Answers to these questions are given in the data which appear in Tables V., VI., VII., VIII., IX., and X.

1. *Complexity of Material.*—All of the observers, as is shown specifically by Tables V. and VIII., gave fewer subjective reports in their descriptions of the 5 than in their descriptions of the 6 figure cards. 66, 86, and 89 per cent. of the reports were given by *W*, *F*, and *D*, respectively, after the exposure of the six-figure cards of Series II. Complexity of material therefore seems to be a condition of subjectivity. This conclusion is borne out by a further analysis of the case records. Table V., under the caption 'Number Added,' shows that more dual additions were given with Series II. than with Series I. After the exposure of cards of Series I., *W*, *F*, and *D* gave respectively 4, 1, and 0 cases of dual additions, whereas they gave, after the exposure of cards of Series II., 16, 15, and 1. The only triple addition reported was by *F*, and this was given after an exposure of a six-figure card (Series II.). These data indicate, therefore, that complexity of material is a condition of subjectivity which is

effective in two ways: (1) by an increase in the total number of additions; and (2) by an increase in the number of dual or multiple additions.[1]

Table V

Showing a Division and Classification of the Subjective Reports Accordingly as They Were Given after the Exposure of Cards of Series I. and of Series II., and According to the Number Given, to the Number Added, to the Position of Those Added; also Showing the Position of the Fixation, the Order of Report, and the Duration of the Exposure in the Experiments in Which Subjective Reports Were Given

Observer.	Series.	Subjective Reports.		Number Added.			Position of Subjective Reports.				Fixation during Exposure.				Order of Report.				Duration of Exposure.					
		No.	%	1	2	3	Not Given.	Left.	Center.	Right.	Not Given.	Left.	Center.	Right.	Not Given.	From Left.	From Center.	From Right.	*a*	*b*	*c*	*d*	*e*	*f*
W...	I.....	37	33	29	4	0	12	19	4	2					3	4	12	14	4	8	6	8	5	2
	II.....	74	66	42	16	0	10	42	14	8					6	6	27	19	11	1	09	9	10	9
	Total .	111		71	20	0	22	61	18	10					9	10	39	33	15	18	15	17	15	11
F...	I.....	13	14	11	1	0	1	0	0	12	4	0	7	1	0	6	5	1	3	3	1	2	1	2
	II.....	80	86	47	15	1	6	3	1	70	17	1	44	1	12	5	45	1	12	9	8	11	14	9
	Total .	93		58	16	1	7	3	1	82	21	1	51	2	2	11	50	2	15	12	9	13	15	11
D...	I.....	3	11	3	0	0	0	1	0	2	2	0	1	0	0	2	1	0	1	0	0	1	0	1
	II.....	25	89	23	1	0	0	5	3	17	9	2	13	0	0	17	6	1	4	4	3	4	6	3
	Total .	28		26	1	0	0	6	3	19	11	2	14	0	0	19	7	1	5	4	3	5	6	4

2. *Position of the Subjective Reports.*—The figures given in Table V. under the caption 'Position of the Subjective Reports' show the number of times that position is not clearly indicated, and the number of times that subjective reports are given at the left, the center, and at the right of the perceptual field.

These data show a 'position of subjectivity'—a position which, for reasons we shall presently consider, is particularly

[1] What is the exact relation, and whether the number of subjective reports varies directly or according to the logarithm of some constant with the number of objective stimuli, are questions which can not now be answered. That some relation does exist, that the number reported is in some way correlated with the number exposed, is quite evident from our results. We hope to return to these questions and to make them the object of a separate study.

favorable for the perception of additional areas. This position is at the left for *W* and at the right for *F* and *D*. It is not affected by the complexity of the exposure material, for the same conditions obtain in the reports of Series I. as of Series II.

The factors which in this experiment condition the position of the subjective reports are:

(*a*) *Fixation.*—The position of the fixation was thought to have been stabilized by the fixation-point of the apparatus. Consequently the observers were not directed to report fixation. Two of the observers, *F* and *D*, fortunately however, gave a report of it in a large per cent. of the experiments.[1] The data for this appear in Table V. under the caption 'Fixation during Exposure.' These data show that where fixation was reported it occurred at the center of the field for both of the observers in about 90 per cent. of the experiments.

A correlation of the point of fixation with the position of the subjective report is given in Table VI. The correlation shows a negative relation. With but 2 exceptions for each observer, subjective reports do not occur at the position of fixation. For example: when *F*'s fixation is at the left no subjective reports are given at the left, though 2 are reported at the right; when his fixation is at the center, no cases are reported at the center, though 2 are reported at the left and

TABLE VI

SHOWING THE CORRELATION BETWEEN THE POINT OF FIXATION AND THE POSITION OF THE SUBJECTIVE REPORT

Observer	Position of Subjective Report	Point of Fixation				Total
		Not Reported	Left	Center	Right	
F	Not reported	2	0	5	0	7
	Left	1	0	2	0	3
	Center	0	0	0	1	1
	Right	25	2	53	2	82
	Total	28	2	60	3	93
D	Not reported	0	0	0	0	0
	Left	2	0	4	0	6
	Center	1	0	2	0	3
	Right	8	2	9	0	19
	Total	11	2	15	0	28

[1] *Op. cit.*, 206, 216.

53 at the right; when his fixation is at the right, however, there is an exception to this rule, for 2 of the 3 cases given here are reported as at the right, while the third is reported as at the center.

These exceptions and the concentration of the subjective cases at the right for *F* and *D* show the influence of other factors.

(*b*) *Clearness.*—Attributive clearness conditions subjectivity in the same manner as fixation conditions it; namely, by acting negatively. The table which gives the position of the upper level of clearness in those experiments in which the dual formation is reported, and the table which gives the correlation between the upper level and the 'position of subjectivity,' reveal that no subjective reports were given in the upper level; that, conversely stated, a one-to-one correlation exists between the position of the subjective report and the position of the lower clearness level. These tables have been omitted from this paper, since the facts are given in Table I. above, which shows that subjective reports are given—when the dual formation is reported—only at the lower level. Attributive clearness therefore sufficiently accounts for the exceptions noted above, and for the concentration of the subjective cases at the right in so far as the subjective reports are given in the dual formation; but this explanation does not of course apply to those cases in which all of the processes are reported as of one level of clearness.

(*c*) *Order of Report.*—The order of reporting we think accounts for the 'position of subjectivity,' in so far as it is not conditioned by fixation and clearness; for, as we stated in our earlier study,[1] and as is shown in Table V. above, all three observers were consciously disposed to a definite order of report: *W* was set to report from the right; *F* from the center, first toward the left and then toward the right; and *D* was set to report as in reading from the left to the right; and the subjective reports, as we have shown above, are grouped at those areas that the observers came to last in their descriptions.

[1] *Op. cit.*, 218.

Table VII. shows the correlation between the order of reporting and the position of the subjective reports.

Table VII

Showing the Correlation between the Order of Report and the Position of the Subjective Report

Observer	Position of Subjective Report	Order of Report				Total
		Not Reported	From Left	From Center	From Right	
W	Not reported	1	3	8	10	22
	Left	4	4	31	22	61
	Center	5	5		8	18
	Right	1	1	5	3	10
	Total	11	13	44	43	111
F	Not reported		2	5		7
	Left	1		2		3
	Center				1	1
	Right	16	13	51	2	82
	Total	17	15	58	3	93
D	Not reported					0
	Left		1	4	1	6
	Center		3			3
	Right		15	3	1	19
	Total		19	7	2	28

This table shows that the subjective reports seldom occur in the areas which were first described. For example: when *W* began his report at the left, only 4 out of 13 subjective cases were reported at the left; when he began them at the center, not one out of 44 cases was reported at the center; and when he began his reports at the right, only 3 out of 43 cases were reported at the right. Similarly for *F* and *D*: *F* began his reports from the center 58 times, a subjective report was given at the center but once; he began his reports from the right but twice, 82 subjective reports were given at the right. *D* began his reports from the left 19 times, subjective reports were given at the left 6 times; he began his reports at the right twice, 19 subjective reports were given at the right.

It might be argued that the order of reporting, instead of being a condition of the 'position of subjectivity,' is itself conditioned by the position of the subjective report. This objection might be pertinent were it not for the fact that the 'Order of Report' was the same for the 'objective' experiments (experiments in which no subjective report occurred) as for the 'subjective' experiments.[1] That being the case, 'position of subjectivity' could hardly act as a determinant of the order of report. Hence we feel justified in regarding the general tendency to name in a certain direction from certain fixed points as a factor which conditions the position at which the subjective cases are reported.

3. *Duration of Exposure.*—The relation of the subjective reports to the time of exposure is shown in Table V. under the caption 'Duration of Exposure.' The letters at the top of the table denote the exposure times employed.[2] These times varied by intervals of 40σ; extending from 40σ the shortest—exposure *a*, to 240σ the longest—exposure *f*.

These data show that the exposure times have no effect upon subjectivity. Subjective reports are given in the descriptions which follow long exposures as often as in the descriptions which follow medium and short exposures. The differences are too slight to be significant, nor are they constant or in the same direction.

4. *Quality.*—It was thought that quality would affect both the number and the position of the subjective reports. We had come to this conclusion from casual observation, for we had frequently noticed in running through the introspections that subjective reports were given in the experiments in which tints and colors of poor saturation were used. If this were the case, if certain colors or color combinations were effective subjectively, we should then expect to find that subjective reports were always given when certain stimulus cards, and never given when other cards, were used.

Table VIII. shows the number of times subjective reports were given for the various stimulus cards.

These data reveal how groundless was our opinion that certain qualities or certain combinations of qualities con-

[1] *Op. cit.*,

[2] *Cf. op. cit.*, 186.

ditioned subjective reports. The occurrence is too evenly distributed throughout the list of stimulus cards to be of any significance. Such variations as occur in Table VIII. must be regarded as due to chance only.

Table VIII

Showing the Number of Times the Subjective Reports Were Given for the Various Stimulus Cards

Series	Observer	Card Number																				Total
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
I......	*W*..........	1	1	1		2	2	2	2	1	1	4	2	2	2	2	1	3		2	2	33
	F..........	1	1		1	2					1	1			1	2		1		1		12
	D..........					1						1	1									3
	Total.....	2	2	1	1	5	2	2	2	1	2	6	3	2	3	4	1	4	0	3	2	48
II.....	*W*..........	1	4	3	5	4	5	5	3	4	3	4	2	1	2	2	1		2	4	3	58
	F..........	3	3		4	4	5	4	4	3	2	4	4	1	2	2	5	3	4	3	3	63
	D..........	2	1	2	1		1	2	1	2	2	1	2	1	1	1	2	1		1		24
	Total.....	6	8	5	10	8	11	11	8	9	7	9	8	3	5	5	8	4	6	8	6	145

Note.—For the specific qualities used and for the selection and arrangement of them on the exposure cards see the earlier study (the footnote to p. 184, and the Table on p. 185).

5. *Practice.*—The effect of practice in the apprehension and description of the exposure cards upon the subjective reports is shown in Table IX.

Table IX

Showing the Number of Experiments in which Subjective Reports Were Given During the Different Parts of the Research

Observer	Experiments			Total
	1–80	81–160	161–240	
W......................	29	32	30	91
F......................	23	27	25	75
D......................	6	7	14	27

These data show that the subjective reports for *W* and *F* were distributed evenly throughout the course of the experiment. Practice therefore had no effect for these observers. *D*'s data, on the contrary, show a gradual increase in the number of experiments in which subjective reports were given. They occur in 6 of the first 80 experiments, in 7 of the second 80, and in 14 of the last 80 experiments. The number of cases is so few and the absolute difference in the distribution is so small that we are, particularly in view of *W*'s and *F*'s results, inclined to the opinion that the variation is due to chance. If this view be rejected, the following fact should be considered before ascribing the distribution to practice: *D* constructed the exposure cards and was thus familiar with the number and quality of the stimuli used. This positive knowledge may have consciously at some times and unconsciously at others limited his descriptions. During the early experiments the effect of this knowledge would be the greatest, and his descriptions would be comparatively free from subjective reports. As time passed, and his memory of the cards faded, and the principles upon which he had constructed them became less clear, his reports would graduaally approach the norm, and subjective reports would appear more and more frequently. From this point of view the increase shown in *D*'s data is due to habituation, to an unprejudiced approach to the problem of describing the exposure card.

6. *Fatigue.*—In order to discover the effect of fatigue the observation hour was divided into three equal parts and the experiments in which subjective reports were given were counted separately in every part. Table X. shows the results of this division and computation.

TABLE X

SHOWING THE NUMBER OF EXPERIMENTS IN WHICH SUBJECTIVE REPORTS WERE GIVEN DURING THE DIFFERENT PARTS OF THE OBSERVATION HOUR

Observer	Third fractional part of the hour			Total
	1	2	3	
W	30	30	31	91
F	25	25	25	75
D	9	10	8	27

The table shows that the subjective reports are distributed evenly for all observers throughout the observation hour and that therefore they are in no way conditioned upon or affected by fatigue.

Summary

The results of this paper may be summarized as follows:

1. The subjective reports were classified upon the perceptual scale obtained in the earlier work. No differences of kind were noted; all the cases fell readily into the cognitive classification. Differences of degree, however, as indicated by the limited range and by the fact that the highest steps were rarely represented, were noted. We are therefore led to the conclusion that 'subjective' reports are as a rule cognitively less clear than 'objective.'

2. The subjective reports were also classified upon the basis of attributive clearness. The same results were obtained as in the earlier work. The experiments were classified into those cases in which clearness was not mentioned—it was not under the conditions of the experiment reportable; into those in which one level was reported; and into those in which two levels were reported. The outstanding difference between the attributive classification of the 'subjective' and 'objective' reports is that no subjective report was given from the upper level of clearness when the dual formation was reported. Our conclusion is, therefore, that subjective reports are less likely than objective to be attributively clear.

3. Since the subjective reports are capable of the same classification as the objective, and since they do not show differences of kind but only differences of degree, they should not be omitted from a perceptual classification upon the ground that they have no known objective correlate.

4. Complexity of material, of all the factors considered, was found to be the only one which conditioned the subjective reports. This was effective in two ways: (1) by an increase in the total number of additions; and (2) by an increase in the number of dual and multiple additions.

5. Under the conditions of this experiment, the duration of the exposure, the quality and the arrangement of the stimuli on the exposure cards, practice, and fatigue have no effect upon the occurrence or the distribution of the subjective reports.

6. Subjective reports occurred more frequently in some areas of the field than in others. The position of this 'area of subjectivity' was different for the observers; it was an individual difference which, however, could be explained and determined by fixation, by the position of the upper level of attributive clearness, and by 'set,' that is, by the general tendency to report in certain directions from certain fixed points.

Journal of Experimental Psychology

VOL. IV, NO. 3 JUNE, 1921

A MIRROR-RECORDER FOR PHOTOGRAPHING THE COMPENSATORY MOVEMENTS OF CLOSED EYES

BY RAYMOND DODGE

Wesleyan University

Any project that involves photographic records of reactive compensatory eye-movements, as in the ocular nystagmus of rotation, is embarrassed by a unique technical difficulty that has hitherto been prohibitive. In order to register eye-movements that are pure compensations, free from visual stimuli and visual controls, the records must be taken either in the dark or from closed eyes. Both conditions make photographic technique somewhat difficult. Photography in the dark seems quite impracticable. Photographing the position of the eye through closed eyelids for a long time seemed no less so.

Neither the dark nor closed eyelids are necessarily prohibitive. But for many years the technical difficulties prevented our completing the description of the five types of eye-movements in the horizontal meridian plane of the field of regard which was begun in 1903.[1] Discussing the reactive compensatory eye-movements in that paper, in contradistinction from the coördinate compensatory, I was obliged to say: "Unfortunately, the essential condition of the phenomenon in pure form is that the eyes be closed. This, however, precludes a quantitative investigation by any of the methods at hand." The first photographic records of the reactive compensatory eye-movements were taken by Holt.[2] He avoided the technical difficulties by frankly

[1] *Am. Jr. Physiology*, 1903, 8, 307–329.

[2] *Harvard Psychological Studies*, Vol. II., 1906, 57–66.

keeping the eyes open, and recorded the eye-movements by our method of the corneal reflection. The difficulty of revolving the long enlarging camera with the subject was avoided by confining his records to the post-rotation nystagmus ('after-nystagmus') which could be photographed by leaving the camera fixed and rotating the subject into position.

From time to time I planned various solutions of the technical difficulties. None of them materialized satisfactorily. Many years ago I tried to work out a pneumatic transmission system depending on the eccentric cornea. Recorders of this order have been used by Schackwitz,[1] Buys[2], and others. The published records show the same faults that deterred me from utilizing this method.[3]

Presumably the best method hitherto published is that of Struycken.[4] He is reported to have photographed a minute polished ball mounted on three feet and attached to the cornea by minute hooks. A total reflection prism in front of the lens permitted simultaneous records of both horizontal and vertical components. I know the method only second-hand.[5] But like the method of the corneal reflection it seems to depend on keeping the eye open. The use of the total reflection prism is a technical expedient of unlimited usefulness. Whether the mode of attaching the polished ball is an improvement over Judd's Chinese white spot[6] and my paraffined tissue paper mounting[7] can only be determined by experience. Personally I confess to some prejudice in favor of a technique that makes the least demands on the subject's ability to stand discomfort.

The possibility of photographing the movements of the eyes in visual darkness by means of the ultra-violet light did not escape us. In view of the possible deleterious effects of ultra-violet light, we hesitated to use it.

In 1916, with the help of my laboratory students we began

[1] *Z. f. Psychol.*, 1912, *63*, 442–453.

[2] *Internat. Zentralbl. f. Ohrenhk.*, 1910, *9*, 57–65.

[3] *Psychological Bulletin*, 1916, *13*, 422.

[4] *Ned. Tijdschr. v. Gen.*, 1918, *1*, 621.

[5] *Zeitsch. f. Psy.*, 1920, *85*, 345.

[6] *Yale Psychological Studies*, New Series, 1905, Vol. I., No. 1.

[7] *Monograph Supplement Psychol. Rev.*, 1907, No. 35.

to record photographically the eye-movements that were hidden under closed lids. This is the method that the present paper is to report.

Theory of the Mirror-Recorder for Closed Eyes

The principle of the recorder is that a surface which presses against the lid and is free to move will tend to assume a position tangential to the underlying corneal surface. That is to say, a mirror which reflects a recording beam of light, will be rotated from side to side as the apex of the cornea passes underneath.

Assuming for convenience an appropriate source of illumination and a suitable recording camera, the relation between eye-movement and records from a mirror resting on the eye-lid over the cornea will depend on the following considerations:

1. A mirror which is pressed lightly against the lid over the cornea will always tend to assume a position tangential to the surface of the cornea on which it rests as far as this is permitted by its mounting. For example, if the apex of the cornea moves clockwise underneath that part of the lid against which a mirror is resting, and if the mirror is free to rotate on an axis that is parallel to the axis of the eye's rotation, as the apex of the cornea passes underneath the mirror it will rotate the latter counter-clockwise. A recording beam of light from the mirror will consequently be rotated counter-clockwise when the eye rotates clockwise. With a fixed source of light its angular displacement will be double that of the mirror.

2. The angular displacement of the mirror will depend on the following factors: (*a*) The angular displacement of the eye. (*b*) The geometrical relationship between the radius of curvature of the cornea and that of the eyeball. The greater the differences in curvature the greater will be the movement of the mirror. (*c*) The relative position of the mirror with respect to the apex of the cornea. The greatest conceivable angular rotation of the mirror would result if the apex of a cornea of zero radius of curvature, infinitely

eccentric to the globe of the eye, just passed the axis of the recording mirror. On the contrary, no movement of mirror would result if it rested on a spherical eyeball while the latter rotated on its geometrical center. (*d*) The thickness and stiffness of the intervening lid. If the lid were totally unyielding all movement of the mirror would cease. If the lid conformed absolutely to the underlying surface it would operate on the geometrical relations as an increased radius of curvature of the cornea. (*e*) The intercurrent movements of the lid. Unless the lid is prevented from moving, voluntarily or by a lid-holder, all lid-movements will operate to move the mirror as though the eye moved.

For obvious reasons it is unnecessary to express these factors in the form of an equation. It would not be simple, and it would be worse than useless in practice if it gave the illusory security in uncontrolled results that mathematical elaboration sometimes produces. A relatively simple geometrical construction (Fig. 1) will indicate the primary relationship between the cornea and globe of the eye as it affects the movement of the recording mirror.

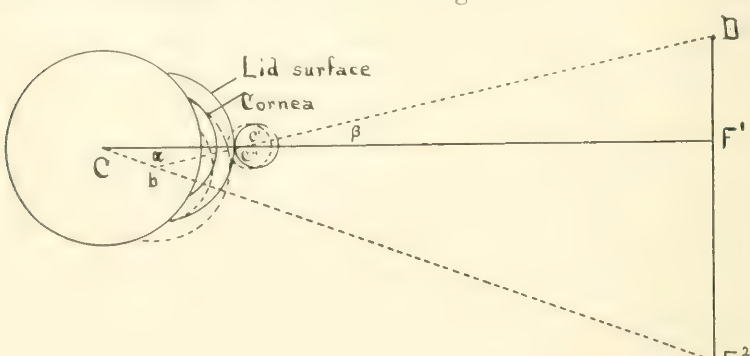


Fig. 1.

In Fig. 1 let CF^1 and the dotted line CF^2 represent successive positions of the line of regard, 20° apart. A recording mirror mounted with its axis on the line CF^1 at the point c^1, and resting on the closed lid over the apex of the cornea, will project a recording beam of light which originates at F^1 along the line of its incidence and back to the point of its origin. When the line of regard moves clockwise to the position CF^2, the mirror will move to a new position tangential to the surface of the cornea which now underlies it, and the recording beam will be deflected counter-clockwise. In this new position of the mirror the line bD drawn from the center of curvature of the cornea at its second position through the axis of rotation of the mirror will be normal to the surface of the mirror. The consequent deflection of a recording beam of light originating at F^1 will be $c''F^1(\tan 2\beta)$. Since $\angle\beta$ is known from our knowledge of the triangle Cbc'', the ideal displacement of the recording light for any angle of eye-movement is directly calculable. Unfortunately, however, we have at present no way of controlling the position of the closed eye with sufficient accuracy to insure the optimum position of the mirror on a principal meridan of the cornea. Our theoretical construction is practically of little use as a basis for the quantitative evaluation of the records.

The thickness of the lid (condition *d*) operates to increase the effective diameter of the cornea about 1.5 mm. to 2 mm.

Movements of the eyelid (condition *e*) seriously complicate every record where they occur. The experimenter must learn to recognize the main forms of lid curves. The most common seems to be due to the lid-reflex and consists of a movement of the closed lid downward and nasalward. This always produces a sharp break in the eye curve in the same direction, with a slow recovery. It is quite unmistakable. Lid-movements which are coördinate with eye-movements and troublesome lid-tremors may be eliminated by a suitable eyelid holder pressing the lashes or a fold of the skin of the eyelid against the orbital edge of the superior maxillary bone. I have found no way to eliminate the reflex lid-movements. A lid-holder only reduces their amplitude.

It is obvious that our warning against trusting the geometrical constructs is quite justified. Each set-up must be empirically controlled, and the empirical data must be checked as often as necessary to prevent changes in the significant factors.

In my own case I have been interested principally with the time relations of the various phases of the eye-movements. Quantitative spatial interpretations will force themselves on our attention in a later paper when we come to discuss the adequacy of the eye-compensations. In that connection we shall discuss the empirical controls. All forms of eye-movement recorders, even under the most favorable techniques with the eyes open, are less well adapted to record positions of the point of regard than they are the direction and duration of movements.[1] In this respect the mirror recorder is no exception. All spatial interpretation of eye-movement records must be undertaken with full understanding of the difficulties involved. They have as much credibility as their controls guarantee, no more.

With this proviso the mirror recorder furnishes a useful and economical instrument for a considerable number of experimental projects and class-room demonstrations. It will demonstrate the main characteristics of the eye-movements in reading, to an entire class. Wherever binocular reading is not demanded it will record the number, duration, and approximate positions of the reading pauses in experimental studies of reading. Its technical advantages over other forms of photographic recorders for these purposes are its ease of adjustment, freedom from abnormal lighting, economy of material and time, and low first cost.

Wherever conjugate movements of the eyes may be postulated, this recorder furnishes the simplest experimental technique for control of the eye-movements in experimental studies of any mental process where they may be supposed to interact. It is adequate for the study of the velocity of eye-movements, and the characteristics of pursuit movements. It is more especially indicated whenever it is desirable to observe or record the action of closed eyes.

Construction of the Mirror-Recorder

The principles of the mirror-recorder permit the widest possible variations in construction to meet experimental conditions. Wherever the head may be fixed we use a simple adjustable arm to hold the mirror against the lid.

[1] *Psychological Review Monograph Supplement,* No. 35, pp. 84 ff.

This would be represented by one of the mirror-arms of the complete instrument which we will describe in detail.

Our mirror-recorder was designed to meet the following requirements:

1. A rigid frame which can be attached to any shaped head and will not vibrate when the head moves. This must be so made that the adjustable attachments for the recording mirrors will interfere least with the supports for the head and other experimental attachments.

2. Attachments to this frame must hold the mirrors against the eyelids with a gentle and even pressure. They must be adjustable horizontally to interocular distance; vertically, to the horizontal meridian of the eye. Secondary adjustments must permit the direction of the recording beam to the slit of the recording camera.

3. The recording mirrors should project a sharply defined image of suitable shape and dimensions. They must be so mounted on the arms that hold them, that they are free to rotate in conformity to the position of the cornea. At least, they must rotate freely on an axis that is parallel to the axis of the eye-movements which they are to record.

Several successive models of the recorder were constructed to meet these requirements. Only the final form will be described. It is pictured in Fig. 2 and in Fig. 3.

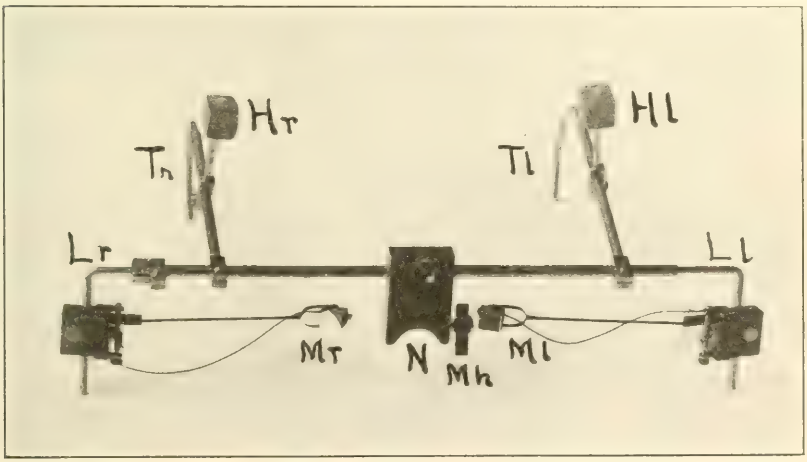


Fig. 2.

The frame resembles a spectacle frame with adjustable nose-piece, temples, and secondary side supports to the head. These five supports make the frame quite rigid and free from accidental vibration. The skeleton of the frame is constructed from $\frac{1}{8}$ in. brass tubing and resembles in form a capital letter *E*. The central offset from the stem would

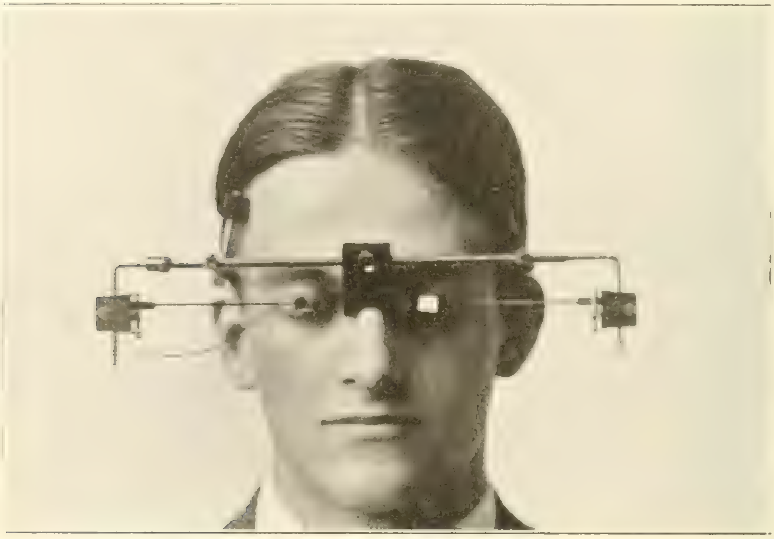


FIG. 3.

represent the nosepiece, Fig. 1, *N*. It is made of hard rubber, shaped to rest on the bridge of the nose, and may be clamped to the main stem at any angle to the temples. The longer offsets at the top and bottom of the *E* represent the temple tubes that hold the adjustable "riding" temple bars *Tr* and *Tl*. The temple tubes are made of $\frac{1}{8}$ in. tubing like the main stem. They may be adjusted to any position on the main stem corresponding to the width of the subject's head. The riding temple bars fit somewhat snugly into the temple tubes and are adjusted to fit over the ears of the subject when the frame is put on.

Into the ends of the main stem tube are thrust the longer arms of two L bars, *Lr* and *Ll*. On their shorter arms these L bars carry hard rubber blocks which may be clamped at various heights and angles. To these are attached the small cone bearings that carry the light steel arms on which the recording mirrors are pivoted, *Mr* and *Ml*.

Starting with the recording mirrors we may recapitulate the supporting members as follows: Each recording mirror is held against the closed eyelid by a light forked, steel bar (see Fig. 3). The axis of this bar is on a hard rubber block which is clamped to the short arm of an L bar. The long arm of the latter slides in and out of the main tube stem to adjust the mirrors to the correct pupillary distance, and to the position of the resting eyes. When properly adjusted these bars hold the mirrors gently against the eyelids. They are free to move horizontally as the apex of the cornea passes underneath the mirrors. The mirror arms terminate in a fork which is drilled to receive a needle, the axis of the small block that holds the mirror.

In adjusting the frame to a subject we follow this general routine:

1. The temple tubes are loosened on the main stem and separated just enough to permit the riding temple bars to clear the subject's head.

2. The riding temple bars are adjusted behind the ears to hold the frame firmly in place without discomfort.

3. The vertical offsets from the temple tubes, *Hr* and *Hl*, are pressed gently but firmly against the head above the frame.

4. The subject is then brought into position on the experimental table, facing the recording camera. The head-line mirror *Mh* is adjusted until the image of the recording light falls across the exposure slit of the camera.

5. Then, one after the other, the mirror arms are adjusted to bring the eye-recording mirrors in contact with the lids at such positions and angles that each rests as accurately as possible over the apex of the cornea, while projecting the recording beam of light onto the exposure slit of the recording camera.

Fig. 3 shows the instrument in position for binocular records.

The recording camera for use with the mirror recorder should accommodate photographic film or paper as wide as five inches. We find the Dodge-Cline camera useful. It takes plateholders which are fitted with adapters for paper.

Our experimentation with sources of the recording light extended over several months. We find the most satisfactory source of the recording beam to be a commercial 100-watt nitrogen-filled incandescent lamp with a horseshoe-shaped filament. This we house to cut out irrelevant illumination, and place on the recording camera at a suitable height, with the plane of the filament in the sagittal plane of the subject's head. A plane mirror placed just above the slit at forty-five degrees to the vertical front of the camera reflects the recording beam to the concave mirrors of the mirror recorder. These in turn project a sharp image of the incandescent filament across the exposure slit of the recording camera. We find the horseshoe filament preferable to the arc light not only because it requires less attention and is steadier, but also because a recording line crossing the exposure slit is safer than a point of light from the arc-crater. Such a line does not leave the slit in slight movements of the head.

Any rapid photographic paper will give satisfactory records if the illumination is properly adjusted. We are best pleased with 'Insurance Bromide No. 2.' For slow records we interpose a filter between the lamp and the eyes. Paper will give excellent records up to 6 cm. per 1″. Beyond that speed, if the slit is kept fine, we get better results with plates.

A METHOD FOR MEASURING RETINAL SENSITIVITY[1]

BY CAPTAIN PERCY W. COBB, M.C. AND MILDRED W. LORING

Medical Research Laboratory, Air Service, U. S. Army, Mitchel Field, Garden City, L. I., N. Y.

ACKNOWLEDGMENT

The authors wish to acknowledge the courtesy of Nela Research Laboratories, for extending the facilities of the laboratory and permitting the expenditure of time, without which this paper could not well have been completed.

INTRODUCTORY

In examining the candidate for flying status during the national emergency the search for a means of rating the visual capacity better and more complete than the usual letter test recalled the fact, already known to workers in sense-physiology, that attempts to measure visual sensitiveness to perception of small contrasts gave highly variable results, ambiguous unless time were taken to accumulate a large number of data, from which the effect of the variations could be eliminated by taking the average. The variations in such results are comparable, in their order of magnitude, to the very quantity which they measure, and the attempt to make such measurements, within the limits of time available and of equipment as yet adapted to the use of the ophthalmologist gave results which were unsatisfactory and meaningless.

A test at one time experimentally practiced in the Air Service involved the use of a smoke wedge.[2] The subject was to draw the wedge slowly across a sight-hole until the

[1] Published with permission of the Surgeon General, U. S. Army, who is not responsible for any opinion expressed or conclusions reached herein.

[2] (*a*) 'Manual of Medical Research Laboratory,' Washington, Government Printing Office, 1918, p. 140, and pp. 145, 146; also (*b*) Air Service Medical, same, same, 1919, p. 270, and pp. 275, 276. The text referred to in these two volumes is almost if not quite identical.

test object ceased to be distinguishable. This latter was a 1-inch square of gray, mounted centrally upon a 2-inch square of lighter gray paper 'where there are 13 perceptible differences between the two squares' (Nos. 32 and 19 respectively, in the series of papers in which white is No. 1, and black is No. 50). As a check, the same device was used, except that the smaller square was replaced by a 20/50 Snellen illiterate 'E' of the same gray. This was the procedure for contrast sensitivity. For threshold sensitivity, the test object used was a 3 millimeter aperture in the iris diaphragm on the De Zeng stand, with specifications as to the lamp used and as to the general conditions of illumination of the room as they would affect the state of adaptation. In all cases the test was to be accomplished in not less than 5, nor more than 8 seconds, in order that the progress of adaptation of the eye during the performance of the test might not be too variable a factor in the result. Average results are in print.[1] Also certain results as to light and color threshold obtained under diminished oxygen during the rebreather test, but these are only stated in percentages based on the total number of cases, in which rise, fall or no change were found; and we are left to guess at the actual number of cases examined, and hence also at the significance of the results stated. The use of the wedge has since been abandoned for reasons which have just been mentioned.

Various letter charts, in which lines of letters are of uniform size, and are in graded degrees of contrast with the background, have also been tried. Notable, among these, is a set included in the work of DeWecker and Masselon which one of us [2] once attempted to use. It soon appeared, however, that there were greater differences of visibility between the various letters of the same line than there were on the whole between the successive differently contrasting lines.

Nevertheless, since the ordinary Snellen chart, or other letter or character chart measures minimal size while presenting maximal contrast, it is eminently adapted to estimating

[1] *Loc. cit.*, (*a*) pp. 288, 289; (*b*) pp. 158, 159.

[2] Cobb, *Psychological Review*, XXVI., 1919, p. 447 ff.

such factors as determine the geometric accuracy of the image, namely, the dioptric conditions of the eye, and it is only in a very incidental way dependent upon the susceptibility of the retina itself to light or light-changes. It apparently does not report changes or differences in retinal sensitiveness within the limits of normal function, and in disease of the retina the impairment in visual acuity is generally enormous. In short, with the test-letters the retina is 'everything or nothing,' and it is this fact that leaves it to be desired that some method be devised which actually measures retinal sensitivity, both as to individual differences, and as to functional differences within the individual.

The selection of the method used in the work herein described arose from considerations developed in a previous communication, to which the reader is referred for a fuller discussion.[1] Briefly, the plan of using a stimulus of minimal area, and of minimal duration, has been adopted and followed as the result of the consideration that these experimental conditions closely represent the common and usual conditions of retinal stimulation occurring in the course of ordinary critical eye-work, and that these conditions are almost identical with those confronting the aviator or aerial navigator who is on the alert to detect an enemy air-craft at the longest possible range. Stated in another way, the method rests upon the use of a stimulus which is within such limits of magnitude, both spatial and temporal, that it may physiologically be considered as being a point and existing for an instant of time; the only implication of this delimitation being the condition that within these limits the product of time and area shall, other things equal, be constant for a given sensory effect. That such a condition may be realized within limits, will appear from what follows.

It is to be added, before proceeding with a description of the apparatus, that this method is perfectly applicable to the study of dark adaptation, with, of course, different limitations involving certain essential changes as to dimensions and as to

[1] Cobb, 'The Momentary Character of Ordinary Visual Stimuli,' *Psychobiology*, II., pp. 237 to 244.

general plan of construction, but without any change whatever as to the principle involved.

But since it is generally, if not universally, accepted that vision at comparatively high intensities of light, or 'day-vision' depends upon organs anatomically distinct from those involved in 'night' or 'twilight vision' and concerned in dark-adaptation, and since it is a fact that the two phases of vision so distinguished have quite different characteristics, it follows that the experimental findings derived from one of these conditions cannot be used as a basis of rating with respect to the other.

Apparatus and Methods

The apparatus used in this work consists of an illuminated screen, *S*, Fig. 2, 76×60 cm., with a small circular aperture at its center. The face of this screen is illuminated by two lamps, as shown in Fig. 1, at the same height as its center, 88 cm. distant from each other and in a horizontal line 50 cm. from the plane of the screen. Back of the central aperture and about 30 cm. distant, is a small screen, S', Fig. 2, illuminated independently by a third lamp, L_3. The three lamps are 6-to 8-volt headlight lamps, vacuum type tungsten filament, the first two rated at $2\frac{1}{2}$ amperes, the third at 2 amperes. The current was supplied by a motor-generator set rated on the generator side, for direct current, at 30 volts, 5 amperes, and at the low voltage used it carried the necessary current (about $6\frac{1}{2}$ amperes) well. The generator being shunt-wound, adjustment was effected by means of a rheostat which directly reduced the current in the field-magnets, and only secondarily by resistance thrown into the external circuit. Slight fluctuations in the current were prevented by floating a battery of seven Edison alkaline storage cells on the line. (Eight cells would have served better.)

When the brightness of the small back screen is just equal to that of the main screen, the observer, sitting directly before the aperture, is not able to distinguish the latter at all, not even at fairly close range. If an opaque object be now interposed behind the aperture, the latter becomes, in appear-

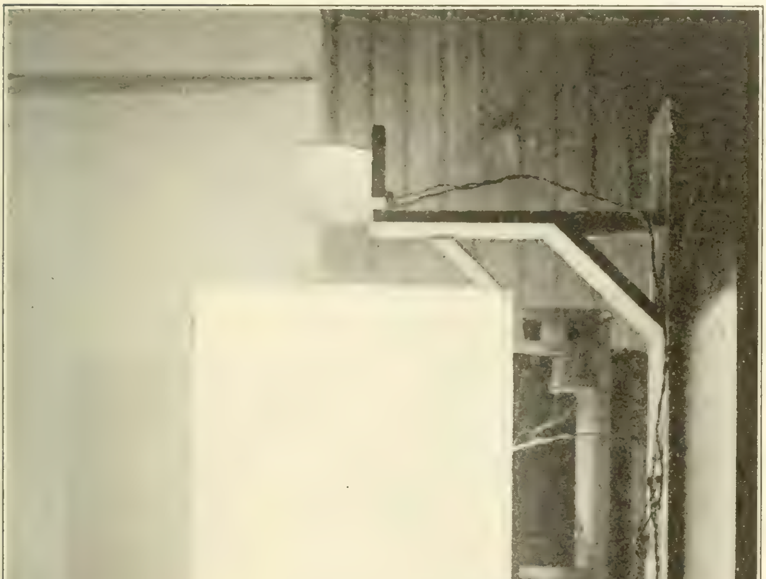


FIG. 1. General view of the apparatus, about as seen by the subject.

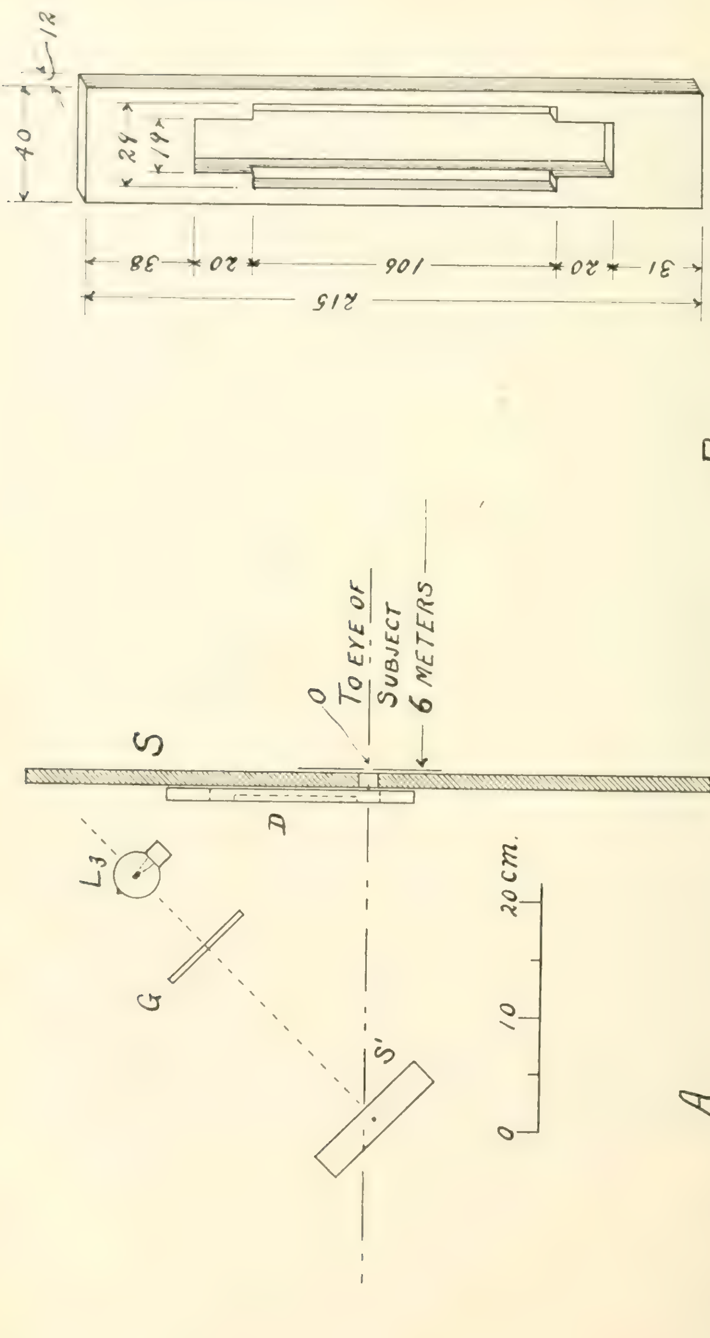


Fig. 2. *A.* Diagram of the apparatus. *B.* Detail of the drop-frame (*D*), with dimensions in millimeters.

ance, a black dot. Experimentally, this condition was to be effected for very brief and measurable intervals of time. A gravity drop-frame was constructed, indicated at *D*, Fig. 2, and arranged so as to fall a fixed distance, in such a way that the aperture was not covered at any part of the drop, and was hence completely invisible to the subject when the frame alone was dropped. A detachable blind could be fitted into the frame to be carried with it in the fall and to pass behind the aperture, cutting off, during the time of passing, the light which had been reaching the eye of the subject from the back screen. This device was so constructed that the upper edge of the blind passed the aperture exactly at the instant when 115 mm. fall, from the starting-point, had been accomplished, the lower edge having covered the aperture a brief interval of time before this. Thus the time of darkening of the aperture depended upon the height of the blind.

The heights of the various blinds were computed to give a geometric series of time intervals corresponding to a series of numerical designations, according to the relation:

$$\text{time} = 200 \times (\tfrac{1}{2})^{N/8}$$

where N is the number of the stimulus and the time is expressed in units of 0.001 second (σ). Thus N, aside from being a convenient numerical designation, has also a definite relation to the time-interval, and fractional values of N are intelligible by means of the formula. The stimuli actually used are included in Table I., where the time in σ, and the height of the blind in mm. are associated with the stimulus numbers. Blinds were constructed for all the numbers, but it was found by trial more expeditious to increase the interval between the stimuli of a series, so the even numbers were laid aside.

Two apertures of different diameters were used. These were punched in the exact center of cards 5 × 8 inches (127 × 203 mm.) and fastened interchangeably at the center of the face of the screen *S*, especial attention being given to see that the aperture was in the correct relation, as to height, with the drop-frame, and laterally with the fixation-point. A hole in the body of the screen was, of course, necessary. This was 11/16 inch (17½ mm.) in diameter.

Table I

The Arbitrary Scale of Stimuli (Stimulus Number) the Duration of Each (Time, σ), and the Vertical Length (Width, Mm.) of the Corresponding Blind, Calculated to Conclude the Required Period of Time Just as the Drop-frame Has Completed 115 Mm. of Free Fall

Stimulus, Number	Time σ	Width, Mm.	Stimulus, Number	Time σ	Width, Mm.
7	109.0	105.4	29	16.2	23.1
9	91.7	96.5	31	13.6	19.6
11	77.1	86.7	33	11.5	16.6
13	64.8	76.7	35	9.6	14.0
15	54.5	67.2	37	8.1	11.9
17	45.8	58.4	39	6.8	10.0
19	38.6	50.6	41	5.7	8.4
21	32.4	43.5	43	4.8	7.1
23	27.3	37.3	45	4.1	6.0
25	22.9	31.9	47	3.4	5.1
27	19.3	27.1	49	2.9	4.3

The diameters of the two apertures were measured from the cards actually used, twice in each of the four diameters at 45°, by the help of a microscope with a vernier micrometer stage and a micrometer eye-piece. The vernier read to 0.1 mm., and the measurements so made checked well with measurements similarly made from a pair of cards punched at the same time with the same punches, but never used. The average diameters were found to be:

	Cards Used	Not Used (Check)
A and *C*......................	4.09 mm.	(4.11 mm.)
B and *D*......................	4.80 mm.	(4.78 mm.)

The corresponding areas are: *A* and *C*, 13.138 sq. mm., *B* and *D*, 18.096 sq. mm.

In order to show whether sensitivity was appreciably modified by minor differences in the brightness of areas about the stimulus area, two slips of sheet metal were so placed that by their shadows the light from each of the pair of lamps was cut off from a portion (something less than half) of the main screen nearest to it as indicated in Fig. 3, leaving a central vertical band unshaded, sufficiently wide to make sure that the margin of the aperture was not touched

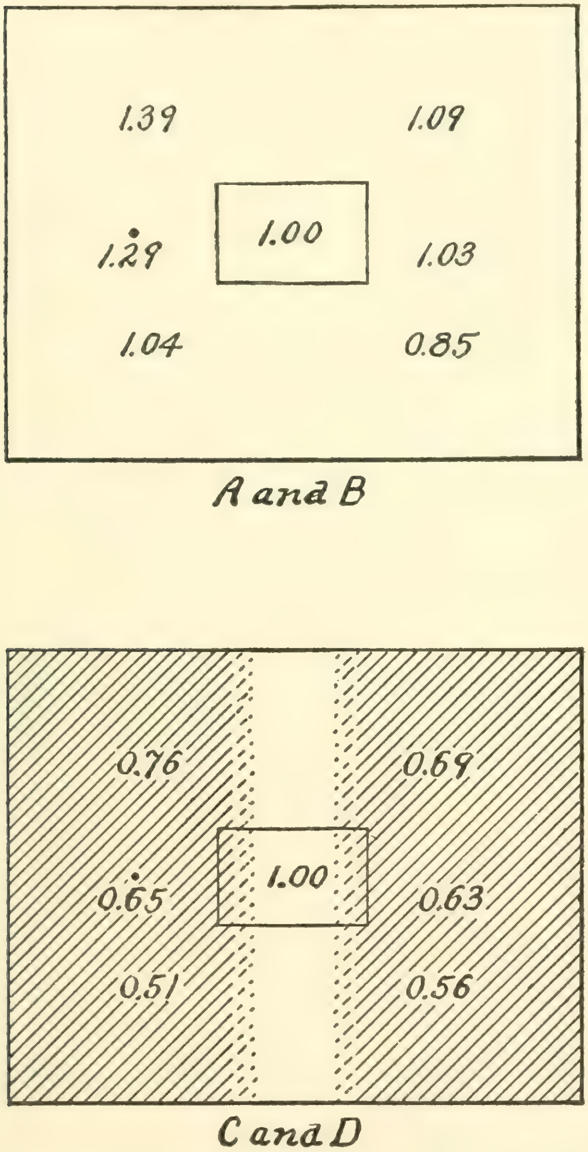


FIG. 3. The relative photometric brightness at various points on the screen, and the distribution of the shadows in cases *C* and *D*.

by the shadows. The change in the distribution of light is numerically shown in the figure, which will indicate that the middle and upper parts of the screen were darkened to the extent of 40 to 60 per cent., the lower part only 30 to 50 per cent. of the central brightness. The shadows were used in the cases under the heads *C* and *D*. The four experimental conditions are thus summarized:

	Aperture	Shadows
A	small	absent
B	large	absent
C	small	present
D	large	present

The screen was viewed by the subject at a distance of 6 meters, the position of his head being fixed by a Troland head rest. Fixation was made upon a point on the screen marked by a black-headed tack, 21 cm. (2° in the visual field) directly to the left of the aperture. The subject released the drop-frame by touching a telegraphic key, conveniently under his hand on the table to which the head rest was clamped. The center of the screen was kept as nearly as possible at a brightness equal to that of the test plate with 3 foot-candles (32.3 meter candles) illumination upon it, which was (the reflection coefficient of the plate being 0.722)

$$\frac{32.3 \times 0.722}{\pi} = 7.42$$

candles per square meter. Readings were taken just before and at the close of each series, and their mean was recorded as the exact value for that series. Photometric measurements were made with a Macbeth illuminometer.

Procedure

The current was adjusted so as to give as nearly as possible the standard photometric brightness at the center of the screen. This was done with the help of the photometer, and the actual reading was recorded. The tendency to a slow and progressive change in the current was found to be minimized by so adjusting the rheostats that the battery was discharging into the line at the rate of two or three tenths of an ampere. The back screen was then adjusted to equal the main screen in brightness, *i.e.*, so that the aperture disappeared for the subject. To accomplish this he sat with one eye in the line of experimental vision, the other closed, and at a distance of about $1\frac{1}{4}$ meters, while the experimenter tilted the back screen until the opening was as nearly as possible invisible. If the experience of the subject were inadequate to make this method expedient, the experimenter performed the adjustment himself, and afterward had the subject pass upon it. The absence of noticeable color difference made this method unobjectionable. The subject then took his place at six meters' distance and the series was begun. At the close of the series a second photometric measurement was made and recorded as before. Four such series, under conditions *A*, *B*, *C*, and *D*, in various orders, were run at one sitting, and occupied about an hour.

The choice of a psycho-physical method to be used in such work as this is largely a matter of excluding theoretically and practically objectionable features, and the discussion of the considerations which led to the choice and adoption of the one here used, and to the present mode of treatment of the results, would be too long to be included in the present communication.

There is comparatively little to be found in the literature upon the method of serial groups,[1] yet in the eyes of one who is interested in psycho-physiology rather than in introspective psychology, this has an overwhelming advantage over other psycho-physical methods in that it provides an immediate practical check on the subject in requiring him to react differentially to the stimuli and to the associated blanks in the same group.

The groups used consisted of 5 equal stimuli and 5 blanks, presented in shuffled order. For the blanks, in order that the motions of the experimenter should not be made the

[1] See Stratton, G. M., *Psychol. Rev.*, IX., 1902, pp. 444 to 447, and Thomson, G. H., *Br. Jour. Psychol.*, V., 1912–13, p. 204 ff. and p. 398 ff. Also Myers, 'A Textbook of Experimental Psychology,' London, 1909, p. 209; same, 2d edition, 1911, Part I., pp. 196 to 197.

basis of reactions, a small horseshoe-shaped metal device was used to replace the blind, fitting into the drop-frame in a similar way, but so formed that at no time during the fall of the drop-frame did any part of it intervene between the aperture and the back screen. A fairly large stimulus was used in the first group, the number of the initial stimulus being later determined by previous experience. The subject had only to fixate the designated mark when the signal 'ready' was given, release the drop by pressing the telegraphic key, and give his answer 'yes' or 'no' as to the adjudged momentary presence of the dot. If upon the first group a perfect score were given, the next group, involving a stimulus of one stage shorter duration, was given without pause; otherwise the next longer stimulus until a perfect score was rendered. The first perfect score being the starting point, consecutive groups were presented with stimuli diminishing step by step until a zero (or negative) score was rendered. This completed the series.

The descending order alone was used. The score for a particular stimulus was simply the number of 'yes' answers given among the five answers to the stimuli, minus the number of 'yes' answers (if any) given among the five answers to the blanks, and could possibly be anything from 5 to -5. Some subjects rarely failed to give correct interpretation to the blanks and rendered a zero score by replying 'no' throughout the group. Others showed greater tendency to fail on the blanks and by the method used all were quantitatively 'docked' therefor, resulting sometimes, but not frequently, is a score of -1 or -2 at the close of the series. The reason for this mode of rating should be obvious: we have no way of knowing whether the subject sees the stimulus except as we may infer it from the fact that his reaction to it is different from his reaction in the absence of it. We can agree with Myers that the positive reactions to the 'catches' are interesting psychologically, but we cannot agree with his statement that they are to be left out of account in the computation. The stimulus has been effective only in so far as the positive reactions to it have exceeded the positive reactions to the blank. The negative score is not troublesome if we remember that it is (so to speak) the obvious consequence of bad luck in guessing, and that in the long run this method of computing causes it and other unlucky guesses algebraically to offset the results of lucky guessing not otherwise detectable.

Estimating the threshold from the serial scores was accomplished by determining, by linear interpolation, that value of stimulus in terms of scale-number which corresponded with a score of $2\frac{1}{2}$ (50 per cent. of a maximum of 5). In cases of inversion, where more than one such point was to be found, the first crossing of the 50 per cent. point was interpolated (*a*) reading the series in the descending direction and (*b*) reading the series in the ascending direction. The mean of (*a*) and (*b*) was then accepted as the threshold for that series.

The performance of four series, involving the four conditions *A*, *B*, *C*, and *D*, occupied about an hour. The order of these was systematically varied from day to day.

The three subjects used are described as follows:

B.—A woman, age about 21, typist. Refraction not determined, but does not wear glasses nor show other evidence of visual defect. An ideal subject, in the fact that she renders prompt and definite judgments.

C.—Captain in the Medical Corps, U. S. Army, specializing in sense physiology, age 47. Astigmatism, well corrected, and moderate presbyopia. Vision, without glasses R. 20/50, L. 20/50 +. Wears: R.: + 1.75 cyl. ax. $92\frac{1}{2}°$, V. = 20/15; L.: −0.25 sph. with + 1.75 cyl. ax. $92\frac{1}{2}°$, V. = 20/15; with + 1.00 sph. each eye, for near vision. As a subject, somewhat slow to react, but with experience in such work.

L.—A woman, age 28, psychologist. Vision 20/20 + 2/8 each eye. Does not wear glasses. Refraction, without mydriatic: R.: + 1.00 sph. with + 0.25 cyl. ax. 82°, V. = 20/15; L.: + 0.50 sph. with + 0.25 cyl. ax. 112°, V. = 20/15. Exophoria, 2 prism diopters at 6 meters. Experienced as a subject in psychological work.

The amount of experience as subject in this identical problem, prior to the work here recorded and incidental to the development of the method, is various for these three. *B* had about the equivalent of 12 series, *C*, about 30 and *L*, about 20. As nearly all of these were run before the method came into its present form, it is only in a rough way that an equivalence can be estimated.

Before computing the results a few of the series were deleted: (3 in the case of *C*, 1 in the case of *L*) where there was suspicion of an error in technic which did not permit of revision, or obvious disturbance of the subject, or, in one case, where the scores of the series gave evidence of some anomalous condition in the subject. In this last it turned out that the inclusion of this value would not have significantly altered the average.

The results of the remaining series were averaged, and in Table II. are given: the mean threshold (N) and its probable error (E_N), and the number of terms (series) entering into the mean threshold (p).

Further, to each group of results (*A*, *B*, *C*, or *D*) Chauvenet's criterion was applied, which resulted in further eliminations (1 in the case of *B*, 3 in the case of *C* and 2 in the case of *L*). The averages from the remaining series are given under (*a*) in Table II.

Results and Discussion

The first question to be examined in the light of these results is the hypothesis implied in one of the introductory paragraphs of this paper. In these results, do the threshold times of exposure and the areas of aperture give equal products for the two cases? If it is true that they do we should find the relation: $a_1t_1 = a_2t_2$, or $a_1/a_2 = t_2/t_1$; in which a_1 and a_2 are the areas of the two apertures and t_1 and t_2 the corresponding times of threshold exposure experimentally found. But it follows from the formula: $t = 200\ (\frac{1}{2})^{N/8}$ that:

$$\frac{t_2}{t_1} = \left(\frac{1}{2}\right)^{(N_2 - N_1)/8}$$

N_1 and N_2 being the two corresponding thresholds as found in terms of the stimulus scale. Hence:

$$\left(\frac{1}{2}\right)^{(N_2-N_1)/8}=\frac{a_1}{a_2}=\frac{18.096}{13.138}=1.3773,$$

Table II

Mean Thresholds in Units of Stimulus Scale

The average results, for each of the three subjects. The small opening has an area of 13.138 sq. mm., the large one 18.096 sq. mm., both seen at 6 meters distance, at the center of a white screen 60 by 76 cm., fixation 2° to the left of the (central) opening.

N.—Average threshold, arbitrary scale. Compare Table I.

E_N.—Probable error of the same.

p.—The number of series upon which the above are based.

Small Opening				Large Opening		
N	E_N	p	Subject	N	E_N	p
A. No shadows				*B.* No shadows		
22.63	0.32	(18)	.....*B*.....	26.05	0.19	(18)
26.31	0.18	(21)	.....*C*.....	29.85	0.37	(21)
22.37	0.29	(20)	.....*L*.....	26.01	0.18	(21)
C. With shadows				*D.* With shadows		
22.79	0.31	(18)	.....*B*.....	27.14	0.21	(18)
26.24	0.24	(21)	.....*C*.....	29.17	0.27	(22)
22.17	0.17	(21)	.....*L*.....	26.96	0.18	(21)

(*a*) Same as Above. Huge Deviations Eliminated by Chauvenet's Criterion

A. No shadows				*B.* No shadows		
22.63	0.32	(18)	.....*B*.....	26.05	0.19	(18)
26.33	0.14	(19)	.....*C*.....	29.85	0.37	(21)
22.63	0.27	(19)	.....*L*.....	26.18	0.17	(20)
C. With shadows				*D.* With shadows		
22.79	0.31	(18)	.....*B*.....	27.32	0.18	(17)
26.05	0.22	(20)	.....*C*.....	29.17	0.27	(22)
22.17	0.17	(21)	.....*L*.....	26.96	0.18	(21)

from which we obtain, by means of logarithms:

$$N_1 - N_2 = 3.695.$$

This is the difference, on the stimulus-number scale, between the time thresholds for the smaller and larger apertures respectively, which is demanded by the hypothesis that the product, area multiplied by time, shall be constant for the threshold.

How far the experimental results go toward confirming the hypothesis from which the foregoing is, a priori, derived, will be seen by a glance at the first three columns of Table III., where the average experimental differences are stated with their respective probable errors. The elimination of the widely deviating results by means of Chauvenet's criterion, it will be seen (*a*, Table III.) has not significantly altered the result. On the average, the difference in scale-units is 3.77, with a probable error $\pm$ 0.14, or (*a*) with application of Chauvenet's criterion, 3.82 $\pm$ 0.14. The deviations from 3.695, the *a priori* value, being respectively 0.075, and 0.125, both lying well within the limits of the probable error. It is to be noted that there are items entering into the final average which depart rather widely from the anticipated result. These are under "*D–C*," Table III., the comparison of the two apertures with the shadows on the screen. The departures of the experimental differences from the value 3.695 are here, for the several observers: *B*, 0.655 $\pm$ 0.37; *C*, $-$ 0.765 $\pm$ 0.36; *L*, 1.095 $\pm$ 0.25, all of which are seen to be larger than their respective probable errors. In connection with this fact, the further fact is to be noted, that the results in the last three columns of the table all indicate that the presence or absence of the shadows was without significant effect, *except* in the cases under *D–B* (shadows compared with no shadows, larger aperture) where the differences are too large, relatively to their probable errors, to be considered insignificant. These differences, further, are all in the same direction as those just pointed out under *D–C*, each for each, and taken together, the two sets of results might be considered to indicate that the three observers differed most widely from each other in the case of the large aperture with shadows on the screen (condition *D*) since this is the experimental condition common to the two sets of differences, each of which, in detail, deviates from the general average in a corresponding way. This can be looked upon as presumptive but not as a definite conclusion, since the differences quoted are not large enough to put the matter beyond doubt.

So far, this discussion has been limited to the consideration of the effect, upon the time value of the threshold, of the two experimental variables purposely introduced. These are:

TABLE III

EXPERIMENTAL DIFFERENCES, IN UNITS OF STIMULUS SCALE

The differences, on the arbitrary scale, due to the use of the large or small opening, and those due to the presence or absence of shadows. Compare Table II.

d.—Experimental difference, arbitrary scale.

E_d.—Probable error of the same.

p.—Weight of the result, the weight of a single series being unity.

Large *vs.* Small Opening				Shadows *vs.* No Shadows		
d	E_d	*p*	Subject	*d*	E_d	*p*
B–A. No shadows				*C–A*. Small aperture		
3.42	0.37	9.00	.....*B*.....	0.16	0.44	9.00
3.54	0.41	10.50	.....*C*.....	−0.07	0.30	10.50
3.64	0.35	10.24	.....*L*.....	−0.20	0.34	10.24
3.54	0.22	29.74	... Mean ...	−0.05	0.21	29.74
D–C. With shadows				*D–B*. Large aperture		
4.35	0.37	9.00	.....*B*.....	1.09	0.28	9.00
2.93	0.36	10.74	.....*C*.....	−0.68	0.46	10.74
4.79	0.25	10.50	.....*L*.....	0.95	0.25	10.50
4.00	0.19	30.24	... Mean ...	0.41	0.20	30.24
All results				All results		
3.77	0.14	59.98	... Mean ...	0.19	0.14	59.98

(*a*) SAME AS ABOVE. HUGE DEVIATIONS ELIMINATED BY CHAUVENET'S CRITERION

B–A. No shadows				*C–A*. Small aperture		
3.42	0.37	9.00	.....*B*.....	0.16	0.44	9.00
3.52	0.39	9.97	.....*C*.....	−0.28	0.25	9.74
3.55	0.32	9.74	.....*L*.....	−0.46	0.32	9.97
3.50	0.21	28.71	... Mean ...	−0.20	0.20	28.71
D–C. With shadows				*D–B*. Large aperture		
4.53	0.35	8.74	.....*B*.....	1.27	0.26	8.74
3.12	0.34	10.47	.....*C*.....	−0.68	0.46	10.74
4.79	0.25	10.50	.....*L*.....	0.78	0.24	10.24
4.12	0.18	29.71	... Mean ...	0.40	0.20	29.72
All results				All results		
3.82	0.14	58.42	... Mean ...	0.10	0.14	58.43

the area of the stimulus-opening; and the presence or absence of shadows upon the screen. In addition to these, there are three other variables, incidental to the experimental procedure, which, including the principal variable, namely the threshold itself, are listed as follows:

1. The time threshold. This is the stimulus number reduced, for the purpose of correlation, to an intercomparable

value by expressing the result of each series in terms of the mean of the group taken under identical conditions. Symbol, *T/M*.

2. The ordinal number of the series in the group of four taken at a sitting.

3. The ordinal of the sitting in the group of 18 to 22 taken under identical conditions.

4. The mean photometric brightness of the field immediately surrounding the test-stimulus.

In what follows these quantities are designated by subscripts corresponding to the numerals above.

These four variables are correlated by Pearson's method,[1] two and two, and from the six correlation coefficients of the zero order, so obtained, the partial coefficients of the second order were computed. These, with their probable errors (E_r), are as follows:

	Subject					
	B		*C*		*L*	
	r	E_r	r	E_r	r	E_r
$r_{12.34}$	.037	.080	.098	.074	−.362	.065
$r_{13.24}$	−.046	.080	.082	.074	.125	.074
$r_{14.23}$	.311	.072	.115	.076	.174	.073

Similarly the standard deviations of the four variables, of the zero and second orders are:

	B		*C*		*L*	
	Zero Order	Second Order	Zero Order	Second Order	Zero Order	Second Order
1	0.065	0.062	0.057	0.056	0.053	0.049
2	1.11	1.07	1.13	1.12	1.12	1.04
3	5.2	5.0	6.4	6.3	6.0	5.9
4	0.12	0.11	0.11	0.11	0.087	0.085

[1] As described by Davenport, C. B., 'Statistical Methods,' 3d ed., New York, p. 44. The mathematics involved in the discussion is treated by Yule, G. U., 'Theory of Statistics,' London, 1919, Chap. 12.

And the characteristic equations derived from these, for the three subjects, are stated as follows:

B:	$x_1 = 0.002\,x_2 - 0.0006\,x_3 + 0.18\,x_4$	$E = \pm\,0.042$
C:	$x_1 = 0.005\,x_2 + 0.0007\,x_3 + 0.06\,x_4$	$\pm\,0.038$
L:	$x_1 = -\,0.017\,x_2 + 0.0010\,x_3 + 0.10\,x_4$	$\pm\,0.033$

In these equations, the various x's have reference to differences or variations in the corresponding variables—or perhaps better, we may say that x represents the deviation of the variable from the mean of all its experimentally measured values, in units the same as those in which the corresponding quantity is expressed.

If we wish to estimate the extent to which x_1 is influenced by the deviation of any one of the other variables, say x_2, from its mean value, we must substitute that deviation of x_2 for x_2 in the equation and put x_3 and x_4 equal to zero. From inspection of the equations, it will thus appear that x_2 is most significant in the case of subject *L*. An idea of its importance in these results may be gained by substituting for x_2 its standard deviation (1.12) and we have $x_1 = -\,0.017 \times 1.12 = -\,0.019$. Which is to say that in this case, fluctuations in the value of *T/M* (x_1) are introduced by the fact that subject *L* gave higher results early in the experimental sessions and vice versa, and that the effect of these fluctuations upon her results are represented by a standard deviation of *T/M* equal to 0.019. This is equivalent to 0.42 or 0.50 in "Stimulus Number" (Table I.) according as we refer it to the threshold obtained with the small opening (average 22.27, Table II.) or the large opening (average 26.48). A representative range of fluctuation due to this cause might be taken as twice this or roughly one number on the scale, if we look upon the fluctuations as centering at the mean, and characteristically extending by an amount equal on the whole to the standard deviation in either direction therefrom. When we investigate in the same way the serial number of the day, the same subject shows the largest deviation: $0.0010 \times 6.0 \times 2 = 0.012$, indicating that the representative practice effect for the whole period of experimentation is small, less than ⅓ of one scale unit. This is, perhaps, not surprising, in view of the fact that each of the three subjects had had practice with the method before any of these recorded results were taken.

The deviations of the photometric value of the test-field, x_4, have not, except in one case, produced any noteworthy disturbance in this work, but such disturbance as is shown, taken together with other considerations, makes the photometric control of the conditions a precaution carefully to be considered. From the characteristic equations, we have, similarly to the foregoing, the disturbance due to photometric fluctuation:

	In Terms of T/M	In Scale-Units
for B, $0.18 \times 0.12 \times 2 =$	0.043	0.97 to 1.14
C, $0.06 \times 0.11 \times 2 =$	0.013	0.34 to 0.38
L, $0.10 \times 0.087 \times 2 =$	0.017	0.38 to 0.45

While, in scale-units, these deviations are not larger than the deviations due to variables 2 and 3, *i.e.*, those defining the place of the series in the experimental order, it must be emphasized, first, that the variations in the photometric values themselves (x_4) are small, being no more than the unavoidable variations in the adjustment of the apparatus to an exact value, but they are still large enough to measure. The largest value of the standard deviation is 0.12 apparent foot-candles (in the cases of B and C) which means that while 3 apparent foot-candles was the brightness aimed at, the actual brightness fell about as often within the limits 2.92 to 3.08 ($3.00 \pm 0.6745 \times 0.12$) as it fell outside of these limits. These limits are the exact value attempted, plus and minus its probable error computed from the actual results.

The second point to be emphasized is, that if this measure of retinal sensitivity is to be used as a test, the apparatus must be checked against a standard light-source, by means of a photometer, and corrected as often as necessary as the work proceeds. The mean coefficient of x_4 in the three characteristic equations is 0.11 and the corresponding coefficients of correlation (14.23) are large enough to be significant. This means that within the present limited range of the actual photometric settings, the measure of sensitivity of the individual changed at the rate of 11 per cent. ($1\frac{1}{2}$ to 3 numbers in the various cases) per apparent foot-candle. While it is not correct to extend the application of this rate of change beyond the range for which it was actually found to exist, it must nevertheless be remembered that the voltage in the usual lighting circuits is often a highly variable quantity; that the percentage change of the candle-power of an electric lamp is several times the corresponding percentage change in voltage; and that an individual lamp operated at its rated voltage is a changing and not a constant source. Such lamps are usually designed and rated to operate 1,000 hours with 20 per cent. drop in candle-power during that period. Twenty per cent. of the present screen brightness is 0.6 apparent foot-candles, which would, on the average, condition a drop of 0.066 in T/M, or a drop of nearly 2 scale-units in retinal sensitivity. And this is nothing as compared with the errors that would have been introduced by depending upon a line-voltage characterized by such fluctuations as those existing at Mitchel Field at the time this work was done.

Another consideration, which led to the installation of the motor-generator and battery used in the present work, was the possibility of errors due to the use of the alternating current. Although ordinarily appearing to give a steady light, electric lamps, and especially those of the lower wattages and therefore more slender filaments, undergo rapid fluctuations in candle-power synchronously with the alternations of the current. On a 60-cycle current the dim (or bright) phases occur at the rate of 120 per second, or at intervals of 8.3σ. Stimulus number 22 has a duration of 29.7σ, number 29, 16.2σ. It seemed clear in advance, that the fluctuations in the light from a lamp on the alternating circuit might introduce errors, variable in amount according to the exact phase of the current in which the shutter was dropped. The use of the direct current evaded this dubious condition; and the use of automobile lamps, designed for low voltage and high current, and consequently having filaments too coarse to readily follow, in temperature, any sudden and slight fluctuations of the current, was especially favorable from this standpoint.

Summary

1. A threshold method of measuring retinal sensitivity is proposed, by which a stimulus of very small and fixed extent is exposed for a very short measured time. The stimulus is physically a negative one, by which is meant that it appears as a black dot upon a white screen.

2. The psycho-physical method of serial groups was used, each series of groups furnishing one datum. The basis of the conclusions is 240 such series, about equally divided among the three subjects.

3. It was found that the product of the time of exposure multiplied by the area of the stimulus was constant, on the average, at the threshold. This applies to the two apertures used, of 13.138 and 18.096 square millimeters' area respectively, observed at 6 meters' distance.

4. For the several observers, the average threshold time, reduced to its equivalent with a standard opening (area = $10\sqrt{2}$ square millimeters) was found to be: *B*, 25.7σ; *C*, 19.5σ; and *L*, 26.2σ.

5. Shadows, covering the major part of the screen, but not the working area immediately about the stimulus, by which 40 to 60 per cent. of the light was cut off, did not significantly alter the time necessary for effective stimulation.

6. The place of the series in the succession of the usual four run at one session did not appear, except in one case, to be of significant effect upon the threshold time. The correlations were dubious in the other two cases. In the one case in question, a manifest hyperopia of one diopter and of one half diopter was demonstrable in the two eyes respectively, and sensitivity appeared to grow less during the progress of the day's series.

7. The serial place of the day in the whole course of experimentation was of negligible import, except possibly in the case of the same observer, where the coefficient of correlation (second order) was 0.125 ± 0.074, and no more than a minor effect upon the result was indicated.

8. There appeared to be a fairly definite and constant correlation between the threshold time and the unavoidable but measurable variations in the photometric brightness of the screen, in the direction of decreased threshold time with increased brightness. This was minimized in the experimental work by careful measurement and adjustment of the brightness. This result justifies the photometric precautions used and indicates the absolute necessity of careful photometric control in any application of this method of retinal sensitivity measurement.

IMPROVEMENT IN BRIGHTNESS DISCRIMINATION AND ITS BEARING ON A BEHAVIORISTIC INTERPRETATION OF PERCEPTION

BY EDWARD S. JONES

Oberlin College

A good deal has been written in support of the necessity of a reaction (*i.e.*, some type of muscular response) in every perceptual experience. The emphasis on reaction can be traced at least back to Dewey in his 'Reflex Arc Concept' article,[1] and is well summarized in Washburn's 'Movement and Mental Imagery.' This tendency, however, has not seeped into many of the current textbook treatments of perception. Is it perhaps because there is little experimental evidence for it, and because of the fact that in such discriminations as those of color or brightness it is hard to explain a difference in perception in terms of reaction?

The following experiment tends to support the belief that even our perception of brightnesses involves a reaction phase; and that the improvement of that reaction is the basis of the improvement in discrimination in that perceptual field. (By improvement in reaction is meant the building up of differentiated reactions.)

The experimental material was a set of shades of gray running from white to black—each shade barely distinguishable from the one next to it. Three observers agreed as to the exact order of sixty of these shades of brightness, and the shades were so numbered on the back from No. 1 (white) to No. 60 (black). This made up the testing series. There was also a 'standard' series of five shades presented before the subject was asked to estimate the numbers of various shades in the testing series. The standard set was made of the shades Nos. 1, 15, 30, 45 and 60, and so numbered below, so that each subject could have a complete survey of the

[1] *Psychological Review*, Vol. 3, 1896.

range of these grays and keep in mind just how to number those seen in the testing series.

The standard was presented each time for 10 seconds, after which 10 of the shades in the testing series were presented in succession, these ten distributed at random throughout the series. The subject merely put down the number which he believed corresponded to the brightness number of each shade of the test series as it was presented. Two such series of ten shades, each introduced by showing the standard, constituted a day's testing for a subject, and he was then allowed to note his errors. Each person's ability to discriminate brightnesses was measured by finding the average deviation of each gray's recorded brightness number from its true number.

Two different groups of subjects were tested, the first to be discussed a control group each subject of which was given twenty shades one day and the same twenty two weeks later. The average deviation of the twelve subjects the first day was 4.8 shades, and later 4.1, showing an improvement of .7 of a shade. These subjects were unselected from an elementary class in Psychology. Eight of the twelve showed improvement, while two remained the same and the other two showed slightly poorer records the second time than the first.

The second group, of ten subjects, were advanced students in a class in educational psychology, several of whom had had a course in experimental psychology. They had been reading Thorndike's surmise that any type of mental function could be improved if interested practise were applied. Further they had received from the writer the suggestion that probably all discriminations depended upon reactions that have been made specific, that we are confused in our perceptions only to the extent that we have not developed reactions which are definite one from another.

This group was given the same series of twenty shades twice, as in the case of the control group, but in between the subjects were given four days of practise. They were also asked at the beginning and at the end of the testing to record any devices they were using as aids in their discriminations, and any muscle strains that appeared or other adjustments which might be of influence. Care was taken to give an entirely different series of grays for the various practise days than that used for the testing of improvement, so that the order of arrangement could not be detected by the subjects.

The results of ten subjects tested accordingly are as follows: Average deviation of initial series of twenty shades (*i.e.*, deviation of estimated from true numbers), 4.0 shades. Average deviation of same twenty shades after five days practise, 2.7 shades. This is of course a marked improvement, as the later figure must be near the average physiological limit for such a function.

The table below records the deviations made by subjects before and after practise, and it also summarizes briefly items given in the introspective reports of subjects. Only one subject did not improve, a girl with an unusually low score at the beginning and who reported clear visual imagery. This subject said she seemed to have less and less confidence in her judgment as she went on.

In the case of 6 of the 7 subjects who improved considerably, one shade or more, definite muscular strains were observed. The other subject of this group reported no variety of muscular reaction, though his final record was superior to that of the others. However when questioned further he replied that he had developed a 'system' which seemed to work. This was to bring his eyes to rest each time before glancing at a shade (so far as any definite fixation was concerned), then to make a number of quick glances, closing his eyes before each. It is quite probable therefore that his mechanism of reaction involved muscles of accommodation or convergence.

Associations, another method of reaction, were apparently useful (*i.e.*, calling a specific gray slate color, or thinking of a black as belonging a certain distance down a well).

Clear visual imagery of the standard was effective, but perhaps not as significant as would be thought. The clearness of this imagery was not enhanced as practise went on and as discrimination improved. The suggestion is clear from the reports that a visual image of the standard is really in terms of a reaction; that perhaps all imagery is fundamentally a mode of reaction. What was frequently referred to as 'holding the standard in mind' is really maintaining certain muscular adjustments, usually of the eye muscles, so that they can be later compared with other reactions. The muscle strains used in connection with these visual images may be the only ones used, though it is quite possible to work up other 'systems' of reaction (such as general body strain) to hold or recall the position of a shade in a series.

Practise with such a system involves the same situation as is present with any form of motor learning. There are random excursions and a more or less unconscious gradual adoption of a technique. Confidence of success is low and unrelated to actual success. *E.g.*, the two subjects 8 and 9 (in the table) said they gained in confidence considerably, while the subjects 2 and 4 claimed to have developed very little confidence even at the end.

It is probable that practise in emphasizing a system, or method of reaction, is valuable in improving such a perceptual discrimination. As in the case of all learning, improvement is better if we are conscious of making correct reactions. The reactions incident to ordinary visual imagery are on the other hand of long standing, and usually beyond the scope of consciousness. Practise may have little influence here, except as some other system is developed to corroborate or confuse the visual image.

The perceptions of brightnesses, or of colors (for similar experiments have been carried out with colors), seem to resolve fundamentally into methods of reacting. With colors the reaction is more apt to be a verbal one, the name of the color; and the discrimination of unfamiliar tints therefore takes a longer time. But without some muscular adjustment, or tightening, there would be no machinery for the discrimination of one stimulus from another. Muscles are the organs of movement and change. They retain habits and images. Nerves are perhaps connectors and conductors only.

Individuals Arranged in Order of their Improvement

Subject	Ave Dev. 1st 20 Times	Ave Dev. after 5 Practises	Introspections as to Imagery and Types of Reaction Experienced
1	5.5	2.2	Noticed eye and general body strains. Used associations (medium gray = slate). Fair visual image of standard without improvement after practise.
2	5.5	2.8	Clear visual image of standard from beginning. Eye movement noticed. All body muscles tend to tighten with light shades.
3	3.8	2.1	Visual image fair, but faded quickly. Eye muscles seemed to tighten with dark shades.
4	3.7	2.0	Visual image was only fair and became no better. Names were used for different shades. Some eye strain observed.
5	3.1	1.6	Visual image of the standard, good. No general body strain noticed. Rapid glancing at stimulus of value.
6	4.0	2.7	Visual image, fair. General body strain as shades were darker. Some associations used.
7	3.9	2.8	Visual image very poor. Eye strains noticed. Rigid fixation on the stimulus helped.
8	3.7	2.9	Visual image, fair. Some eye and forehead strain noticed as shades were darker. Not always present however.
9	3.9	3.2	Visual image, very poor. Some tendency to tighten with dark grays. Otherwise no strains.
10	2.5	4.3	Used visual image entirely. Very clear, but seemed to lose confidence in it. No strains observed.

PHYSICAL ATTRACTIVENESS AND REPULSIVENESS

BY F. A. C. PERRIN

University of Texas

A study of the physical characteristics of human beings has its place in psychology, as well as in the various interests represented by art, anthropology, anatomy and physiology. If we may include under this heading, not only static and anatomical items, but certain forms of expressive behavior, personal habits and mannerisms, and qualities of the voice, we will find additional justification for the claim just made. The published accounts dealing even indirectly with this group of factors are extremely meager, both in scope and in amount. Until the recent appearance of Dunlap's *Personal Beauty and Racial Betterment*, they were to be found, only after diligent search, in the writings of Ellis, Thomas, Freud, and a few others. It is the purpose of the present paper to suggest a few theoretical considerations, and to submit the results of a statistical investigation, both concerned with this general subject.

The physical characteristics of an individual may be studied from two distinct points of view. They may, in the first place, be considered with reference to their possible diagnostic values. While this does not fall within the scope of the present investigation, it deserves passing mention. In spite of the wild claims of the professional character analyst, it is possible that physiognomy errs more in its violation of scientific method than it does in its major premise. The psychologist is insistently a seeker after signs; and if they can be found in delayed verbal responses, psychogalvanic reflexes, and systolic manifestations of blood pressure, they may presumably be discovered in those observable aspects of behavior which are intimately associated with emotional and intellectual life. In particular, some of the items enumerated in a subsequent section of this paper, under the heading of "Expressive Behavior," might well repay serious investigation from this point of view.

Of more direct importance, however, is the fact that the physical characteristics of an individual constitute a large group of the total series of effects produced by that individual upon others.[1] The question is whether or not the total series of effects, or any part of that series, is deserving of psychological attention. And it is, if we are to assume that the stimuli which elicit responses in the human animal are at all worthy of scientific interest. For the human animal responds primarily to social situations; his conditioned reflexes center around people, as well as physical objects; his intellect and skill, even, are largely socialized. Aside from this general consideration is the fact that physical characteristics have their significance for vocational and for individual psychology, and for a science yet to be formulated—the psychology of personality and character.

The statement that people do react definitely to the physical in other people may be made safely without statistical evidence. Most of us respond unambiguously to physical beauty and ugliness, to a pretty face or a well-proportioned body. And these factors are not always made subordinate to intellectual and moral qualifications. Even in Y. W. C. A. posters, we find occasionally that delineations of feminine physical attractiveness prevail over suggestions of goodness and virtue. Just why the physical characteristics of individuals should exert so profound an influence over their associates furnishes an interesting topic of speculation. A list of the factors involved would certainly include sex appeal, and the muscular responses elicited by the manifestations of bodily symmetry, proportion and balance, bodily rhythm and physical strength, grace and awkwardness. The evidence offered in this paper seems to favor sex appeal as the fundamental factor. But whatever the explanation may be, the fact of physical appeal remains as an important aspect of human behavior.

[1] For a discussion of the general point involved, see the beginning chapter of Thorndike's *Individuality*.

With the above considerations in mind, the writer undertook a number of statistical investigations in the psychological laboratory at the University of Texas. He was most ably assisted by a group of advanced students in the department at that institution. These students were well qualified for the work in one essential respect: they were acquainted socially with the large number of students who unconsciously acted as subjects for the investigation. The results here reported represent a survey of the field, rather than an intensive study of any one phase of the subject.

1. *A Rating Chart for Physical Characteristics.*—The first aim of the writer was to obtain an inventory of the elements which operate in positive or negative physical appeal, and this was realized in the formulation of a rating chart for physical characteristics. No special merit, of course, is to be found in a chart of this nature. In itself, it is merely an itemized list of the factors involved, and its value depends upon the purpose for which it is used. The items here enumerated, however, have a certain intrinsic value, inasmuch as they were determined by an actual inductive procedure. The members of the experimenting group were asked to list 25 girls and 25 boys noted for their physical attractiveness, and to list a corresponding group noted for their negative appeal. The subjects were selected from the entire undergraduate body, and they represented, respectively, the extreme cases in the entire range. With the names before them, the judges were instructed to write out all possible characteristics which aroused affective reactions of any sort. These characteristics were then classified, somewhat arbitrarily, and were printed on a four-page folder, in tabular form. Seven columns were provided for check marks, indicating the ratings to be made on any one individual regarding the amount or predominance of the trait or characteristic, and a corresponding set of columns was provided for the ratings on the degree of pleasing appeal of each characteristic. The columns were numbered, and the numbers were defined as follows: 1, Very Low; 2, Low; 3, Slightly Below Average; 4, Average; 5, Slightly Above Average; 6, High; 7, Very High. Complete directions were

given for scoring. The items as classified, with many detailed explanations omitted, are given in the following table:

A. STATIC PHYSICAL CHARACTERISTICS: BEAUTY, SIZE, PROPORTION, SYMMETRY, TEXTURE OF HAIR AND SKIN, ETC.

I. *Head and Face*

Size of head
Shape of head
Symmetry of features
Heaviness—lower part of face
Skin of face—smoothness
Skin of face—color
Skin of face—freckles
Skin of face—tan
Hair—amount on pate
Hair—amount on face
Hair—darkness
Hair—amount of red
Hair—glossiness
Eyes—size
Eyes—color
Eyes—distance apart
Eyes—large sockets
Eyes—horizontal position
Eyes—puffiness below
Eyes—wrinkles around
Eyes—dark below
Eyes—expression
Eyebrows—amount
Eyebrows—distance apart
Eyebrows—straight
Eye lashes—length
Ears—size
Ears—protruding
Nose—dilate nostrils
Nose—size
Nose—shape
Nose—correct color
Teeth—white
Teeth—regular
Teeth—protruding
Teeth—gap-toothed
Mouth and lips—size
Mouth and lips—color
Mouth and lips—smooth
Mouth and lips—symmetrical
Mouth and lips—same size
Mouth and lips—expression
Chin—size
Chin—shape and position
Chin—dimples
Cheeks—full
Cheeks—wrinkles
Cheeks—dimples
Cheek bones prominent
Forehead—width
Forehead—height
Forehead—shape
Neck—thickness
Neck—length
Neck—uniform diameter
Face—shape

II. *Arms and Hands*

Arms—size
Arms—shape
Arms—length
Elbows—pointed
Elbows—size
Elbows—dimples
Elbows—red
Elbows—rusty
Wrists—size
Wrists—fat
Hands—size
Hands—length
Hands—width
Hands—big joints
Hands—nature and shape
Hands—moist
Hands—prominent veins
Nails—beauty

III. *Legs and Feet*

Legs—size
Legs—length
Legs—shape
Legs—bowlegged
Legs—knock-kneed
Ankles—size
Ankles—beauty
Feet—size
Feet—length
Feet—width
Feet—beauty

IV. *Trunk*

Size
General proportions of trunk
Proportions of bust
Proportions of hips
Proportions of waist
Proportions of shoulders
Width of shoulders
Slope of shoulders

B. Personal Habits

Care of hair and scalp
Hair arranged in becoming fashion
Hair on back of neck
Hair arranged in prevalent style
Hair on back of neck
Use of perfume on hair
Attention given to facial hair—eyebrows, lashes, etc.
Care of nails
Cleanliness of teeth
Care taken to avoid unpleasant breath
Care in coughing, etc.
Care of eyes—dirt, etc.
Becomingness—care of glasses
Care of skin—cosmetics, etc.
General care of body
Appearance of health

C. Expressive Behavior

General physical pose
Erect carriage
Air of dignity
Suggestion of social training
Vivacity, energy
Nervousness
Pleasing manner of walking
Pleasing manner of standing, etc.
Graceful in use of hands
Graceful in use of legs and feet
Tendency to laugh
Tendency to frown
Tendency to smile
Nature of laugh
Tendency to cry
Tendency to pout
Tendency to show teeth
Use of muscles around eyes
Tendency to use facial expressions
Attitude towards others—external evidences of interest
Tendency to touch other people
Handshake pleasing
Looks at you straight in eye
Attractive street habits
Attractive table manners
Attractive society manners
Attractive home manners
Tendency to use baby talk

D. Voice

Modulation
High pitched
Quality, timbre
Intensity
Distinctness of articulation

E. Dress

Good taste in dress: emphasis upon personal taste rather than upon style; pleasing effects
Tendency to follow prevalent styles
General neatness, orderliness, care

2. *The Extent to Which Physical Elements Contribute to the Appeal of Intimate Associates.*—Having secured an inventory of physical characteristics, it was the next concern of the writer to determine their rôle in the selection made by an individual of his intimate friends and associates. It was thought that they could be isolated and measured only by an experimental situation which would force an individual to rank his intimate associates in the order of his general preferance for them, and then to rank them with reference to a number of abstracted traits, in the hopes that the correlations existing between the order of preference and the various traits would explain the preferences.

This was done with a mixed group of advanced students. It should be noted, of course, that a university student differs in one essential respect from a business or professional man. His associates are not selected primarily on the basis of professional interests. It may be argued, however, that just for this reason the student is the better prepared to analyze the appeal made by his friends. And it is conceivable that the business man temporarily freed from the personal relationships directly due to his vocational environment might show reactions identical with those of the student. But since it was not possible to conduct a control experiment, the following results are offered as characteristic only of the group investigated, with the assumption that they represent a fair sampling of university students.

The members of the experimenting group were asked to write the names of the 10 men and the 10 women for whom they had the greatest degree of personal affection. They were definitely assured that no one would see these names, and they were instructed to represent the individuals listed by the first ten letters of the alphabet. They were then asked, with specific and detailed explanations, to rank these names in their order of personal preference, irrespective of the sex of the individuals listed, and irrespective of extraneous interests of any kind.

They were given no intimation of the nature of the experiment until the lists and rankings had been completed.

They were then given mimeographed forms, provided for the purpose, and instructed to rank the individuals listed with reference to each of a number of characteristics. They were thus forced, by this method, to assign some one in the list to first place, some one to second, etc., as regards each characteristic.

The various rankings were each plotted against the order of personal preference, and the medians drawn. It was thought desirable to establish tendencies, and the relative strength of each tendency, rather than exact correlations. Yule's formula

$$Q = \frac{(AB)(\alpha\beta) - (A\beta)(\alpha\beta)}{(AB)(\alpha\beta) + (A\beta)(\alpha\beta)}$$

seemed sufficient for this purpose, and it was accordingly used. The fact that it disregards differences above and below the median provides against any error due to uncertainty on the part of the judges; for, while they were not always sure of all the relative positions assigned, they were sure that their differences extending over three places were reliable. The following table, then, is submitted as showing only differences in tendencies; and the range of the coefficients seems to indicate actual differences in these tendencies. The various coefficients, each representing a comparison between the order of personal choice and some characteristic or trait, are as follows:

Extent of acquaintance	+.77
Expressive behavior pleasing	+.69
Affectionate disposition	+.69
Individuality of friend	+.66
Sincerity of friend	+.64
Physical attractiveness	+.60
Consistent nature of friend	+.57
Sense of humor	+.57
Accomplishments—social	+.55
Personal habits pleasing	+.55
Physical appeal	+.50
Ethical ideals	+.48
Intelligence	+.48
Voice pleasing	+.38
Vivacity, energy	+.38
Beauty of face	+.37

Optimistic disposition	+.37
Refinement—intellectual	+.37
Beauty—hands and arms, legs and feet	+.22
Dress pleasing	+.13
Beauty of body	+.13

The range of the coefficients does seem to suggest a partial explanation of the basis for personal friendships. Naturally, the friends best liked are the ones best known. And since each judge was instructed to rate on the basis of his own personal reaction, individuality and sincerity are both relatively high. The instructor was warned that the characteristic of sincerity would not necessarily apply to the associates listed with reference to other people. Ethical ideals and intelligence are evidently not considerations of prime importance. A discussion of the results brought out the fact that these presumably higher traits were presupposed only to a certain extent. They were not positive requirements. The fairly low position given to optimism was explained by the fact that intimate friends are likely to be confidential, and correspondingly inclined to talk about their troubles and worries.

As regards the rôle of physical characteristics in the selection of intimate associates, it is evident that static beauty holds a subordinate position, but it is equally obvious that two other groups of physical elements, expressive behavior and affectionate disposition, rank quite high. Expressive behavior was explained and defined in terms of the items given on the rating chart, listed in the preceding section of this paper. These items include many of the elements which ordinarily constitute refinement, and this fact explains the low position given to that trait. Physical appeal was defined somewhat frankly as recognized sex appeal. While some of the judges insisted that it was a factor quite distinct from the appeal made by manifestations of an affectionate disposition, others held that the two appeals represent one and the same thing. It is safe to assume, therefore, that sex appeal should occupy a higher position than the one actually accorded to it. But it was obvious, both in the results of this test and the test reported below, that sex attraction could not be measured by the beauty of the body.

Since the experimenting group consisted of both boys and girls, the experimenter was interested in the possibility of sex differences in the judgments. These differences were not especially in evidence. Of the 21 girls, 13 gave a boy as their choice for first position; of the 8 boys, all gave first place to girls. A further tendency, not at all marked, was discernible for the boys to prefer girls, and for the girls to prefer boys. With the boys, an affectionate disposition was valued when shown by a member of the opposite sex, and sex appeal was predominatly feminine appeal. While the reverse was true, in the main, for the girls, their ratings on the two characteristics made more provision for members of their own sex. No sex differences of any magnitude were to be found in the remaining characteristics.

3. *An Analysis of the Group Characteristics of Physically Attractive and Repulsive People.*—The third investigation was concerned with the physical characteristics of groups distinguished for their positive and for their negative physical appeal. For this purpose, it was thought necessary to study four groups, respectively made up of (1) physically attractive girls, (2) physically attractive boys, (3) physically unattractive girls, (4) physically unattractive boys. The practical difficulties involved in the making up of such lists are obvious. In the case of the first two groups, however, it was decided to face these difficulties squarely.

Seven advanced students, carefully selected by the writer, were asked to write the names of the students on the campus, of both sexes, noted for their extreme physical attractiveness. The names were first written independently, and then compared; and only those candidates who received 5 votes of the 7 were included in the combined lists. The 7 judges then discussed the merits of the candidates, and eliminated a number of them. The writer, of course, had no part in this procedure. Copies of the rating chart were distributed to the judges, with instructions to secure 5 independent ratings on each candidate. Most of the ratings were made by members of the original group of 7; some of them, however, were made by room-mates or fraternity brothers of the candidates, under the immediate supervision of members of the experimenting group. The writer has every reason to believe that the names listed represented the best possible selection, and that the ratings secured were bona fide.

As it was inadvisable to permit a discussion of physically unattractive people, a different method of securing data for the remaining two groups was followed. A number of judges, selected from advanced courses in psychology, were called together, supplied with the rating charts, and instructed to make individual ratings on people noted for their extreme physical unattractiveness. Each judge was asked to make two ratings, for a boy and for a girl respectively. This method made it impossible to avoid the danger of duplication. Small groups of judges did, however, compare names for the purpose of eliminating the duplicates. As an additional precaution, each of the judges was asked to select people who were probably unknown to the other judges. The method has a further disadvantage, inasmuch as it did not permit the use of more than one rating for each subject; but it is probable that a single rating, by a competent judge, when that rating is made on objective traits, is sufficiently reliable for a study of group traits.

The ratings for the individuals in the first two groups were transferred to a new chart, and the median was taken as the final rating for the individual. The ratings for the individuals in each of the four groups were then copied on a larger chart, and the median judgments on the amount of each trait, and the degree of appeal of each trait, for each group, were ascertained. These results were accepted as indications of group characteristics. All questionable or incomplete records were discarded, and the final groups consisted of 27 physically attractive girls, 14 physically attractive boys, 30 physically unattractive girls, 32 physically unattractive boys.

This particular investigation differs from the usual psychological study of groups in one important respect. Ordinarily, a group is distinguished by some common element, shared by each member of the group. In the present case, however, the common element was an effect produced upon others, and an effect due to a variety of causes, no one of which was necessarily manifested by any particular member of the group. Any one element, therefore, might presumably stand a chance of appearing in equal amounts in two opposed groups, unless it clearly belonged to only one of them. In fact, the difficulty of describing group characteristics of this nature is the difficulty which stands in the way of a scientific study of character and personality. But in spite of this, the results did disclose certain group tendencies. They are summarized in the following statements:

(*a*) Anatomical measurements of the attractive people of both sexes conform to the mode or standard. This means that attractive people were given ratings of 3, 4, or 5, predominatly of 4, on those items referring to anatomical characteristics. Such items were those concerned with the size and proportions of parts of the body, and the amount of hair and fat. The degree of appeal made by these characteristics was also of median intensity. Both statements hold for both sexes.

(*b*) The corresponding anatomical measurements for unattractive individuals show a slight tendency to depart from the mode, and the negative affective reactions to these characteristics show a somewhat greater proportionate tendency to exceed the median. In other words, any feature, if it conforms to the mode, tends to elicit a favorable reaction of median strength; but if it departs from the mode, it tends to arouse an unfavorable reaction of greater relative strength. These tendencies hold equally well for both sexes.

(*c*) A few anatomical items disclosed a more definite tendency to depart from the mode, in the case of unattractive individuals. These were found in connection with the proportions and symmetry of parts of the body, and with color. Physically unattractive girls were rated low (in position 2) as regards the color of the mouth and lips, and the proportions of the legs and feet, and hips; physically unattractive boys were rated correspondingly low as regards the proportions of arms and feet. These features, however, were rated only one point above the median (5) in their negative appeal.

(*d*) In terms of the degree of appeal, irrespective of the amount of the trait, the most valuable characteristics shown by attractive girls were clean hair, clean teeth, care taken to avoid unpleasant breath, care in coughing, care of eyes, tendency to arrange hair in prevalent fashion, aristocratic bearing, and general care of the body. These were ranked 6. Two of them, the tendency to avoid unpleasant breath and the care of the body, were also ranked 6 with the attractive boys. The most unpleasant characteristics in unattractive girls were the shape of the chin, proportions of bust, shoulders and hips, the expressions of the eyes and mouth, lack of care of the hair, lack of taste and neatness in dress, absence of aristocratic bearing, poor physical poise. The unpleasant characteristics of physically unattractive boys were quite similar. The eyes, ears, mouth and lips, care of the hands and nails, general care of the body, absence of aristocratic bearing, all came in for unfavorable mention.

(*e*) The results of all the tabulations, including those just referred to, show that physical attractiveness is to be explained primarily in terms of behavior. The items included under the two headings, expressive behavior and personal habits, received more emphasis than the more static items in the list. Beauty is as beauty does. While it is true that a few anatomical items, fairly definitely associated with the sex function, were stressed, these items represented negative, rather than positive, criteria of attractiveness.

A second uniform result is the fact that, in the four sets of measurements, the definite line of separation is to be found between attractive and unattractive individuals. These differ more from each other than the two sexes differ, and the standards of attractiveness are the same for both sexes. This may mean, either that physical appeal is not sex appeal, or that it is sex appeal, with emphasis upon its distinctly homosexual, as well as heterosexual, nature. In the light of the results obtained from the investigation described below, the second interpretation seems to the writer to be somewhat more adequate.

4. *Physical Attractiveness and its Correlated Traits.*—

Physically attractive individuals very naturally elicit fairly constant and uniform responses from other individuals. If we may assume that character is in part a product of social encounter, we may well believe that common traits are to be found in people who stimulate common responses. The psychologist, for instance, who is interested in the question of sex differences, must face the presumption that such differences exist, in spite of the fact that they have been minimized by recent investigations. In these investigations, tests of original nature have been largely employed; but if the conditioned reflex means anything at all, it means that people differ, not so much as regards their instinctive equipments, as they differ regarding the stimuli which bring them forth. So far as social stimuli are concerned, sex distinctions are most certainly made. And distinctions just as great, if not greater, are made between physically attractive and unattractive people. Any study of group differences which stresses innate differences exclusively is necessarily misleading, if it purports to give a picture of actual, existing differences.

In an attempt to secure data upon this question, the writer asked four groups of three judges each to rate a number of their associates, of both sexes, on a number of character traits. All possible care was taken to secure trustworthy results. The median judgment was used as the final rating, and final ratings on 120 undergraduate students, 70 girls and 50 men, were secured. In the treatment of the data, the method of unlike signs again was used. The object was to establish tendencies only, and this method was believed to be sufficient for that purpose.

The traits listed fall into three classifications—namely, general intelligence, attitude towards life, and sex attitudes. The results are as follows:

(*a*) Absence of correlation obtained between general intelligence, as determined by character ratings, and physical attractiveness. The coefficient found was +.02, and the traditional antithesis between beauty and brains thus seems to be without support. In fact, if any a priori case could be made at all, it would favor a positive rather than a negative relationship. Physical attractiveness presupposes a certain amount of good health, and health, in turn, is certainly not antagonistic to mental vigor. Again, physically attractive people are in demand; they are more likely to mingle with people than are unattractive individuals, and as a consequence, to acquire the wisdom which results from such intermingling. While no correlation was attempted between university grades and physical attractiveness, the presumption, again, would be in favor of a slight positive relationship in the group investigated; for the scholastic requirements at the University of Texas are higher for fraternity than for non-fraternity students. Needless to say, the attractive people were recruited predominantly from the fraternity groups. To offset these influences, the fact remains that the attractive student in a coeducational institution is not forced to rely upon mental ability alone, either for academic or for social success. The coefficient found, therefore, probably represents approximately a true relationship.

(*b*) If it is true that a social premium is placed upon physical appeal, it would follow that individuals so endowed tend to live a more active social life than individuals less in demand. As the result of this greater emphasis upon social activity, we should expect to find a greater practical knowledge of human nature. And we should expect, too, that this knowledge would be concerned with sex attitudes. With this possibility in mind, the writer described three general attitudes, and asked for ratings on these attitudes. The three were included in a total list of 33, and the judges had no intimation of the immediate use to be made of the data. An idealistic attitude was defined in terms of conscious emphasis upon ideals—religious, social, ethical, or æsthetic; a sophisticated attitude was explained as a practical knowledge of human nature, particularly as it is manifested in sex relationships, with a reliance upon that knowledge, rather than upon some ideal, as a basis for conduct; and a blasé attitude was described as a sophisticated point of view, plus an acceptance of human nature as it is found, with no tendency to be shocked at some of its weaknesses.

These character estimates were each correlated with a trait described as "Liking for Opposite Sex." This trait was defined from a behavioristic standpoint. It presupposed, therefore, not only a passive liking for the opposite sex, but an active tendency to manifest this disposition.

The coefficient found between the ratings for this trait and the ratings for idealism was −.31, while sophistication yielded a coefficient of +.41. and blasé attitude, +.13. The distribution was approximately the same, on each of the scatter diagrams, for the two sexes. Comment is perhaps superfluous, but it may be stated that the results are about what one should expect.

(*c*) A number of traits were described, all concerned more or less directly with physical attractiveness and sex attitude. The relationships among these traits, as established by the *Q* formula, are as follows:

General physical attractiveness and appeal to opposite sex	+.87
General physical attractiveness and appeal to same sex	+.52
Appeal to opposite sex and appeal to same sex	+.41
Liking for opposite sex and liking for same sex	+.35
Appeal to opposite sex and liking for opposite sex	+.58
Energy, vivacity, and appeal to opposite sex	+.44
Energy, vivacity, and appeal to same sex	+.38
General social ability and physical attractiveness	+.71
Good taste in dress and physical attractiveness	+.83

It is evident that traits closely associated with positive sex appeal tend to show positive correlations. It is also obvious that our results do not indicate the presence of a line of demarcation between sex appeal and physical attractiveness. Which is the larger and more fundamental term of the two is a matter of conjecture, or possibly, of definition; but, all results considered, one may entertain the hypothesis that the traits listed represent a linear series of responses, with overt sex desire at the maximum end of the series. Sex appeal exists, and since it can not be readily isolated from other appeals, we can only assume that it operates through them, and that it is to be identified with them. If fondling and physical contact—those aspects of behavior which constitute an affectionate disposition—are sexual in nature, we may find added support for this interpretation in the high place given to that trait in the investigation described above.

APPARATUS FOR THE STUDY OF VISUAL AFTER-IMAGES

DONALD A. LAIRD

University of Iowa

As a rule the experiments with positive visual after-images are among the least successful of the experiments in the introductory laboratory work in psychology. The simple and easily constructed arrangement to be described has been found to be very effective in demonstrating, under a variety of conditions, this phenomenon which is usually rather elusive to the beginning student. Some research value attaches to the apparatus due to the simultaneous arousal of both the positive and negative after-images under comparative conditions and the observations on the course of retinal adaptation or fatigue which can be approached by way of after-images with the device to be described. The simultaneous appearance of the two images favors the identification of the positive and negative by the uninitiated. The sequence of temporal changes in the images is especially striking and pertinent to the functioning of the retina. The beauty and uniqueness of the phemonema elicted with this arrangement insures an active interest on the part of the observers.

The apparatus consists essentially of a reflecting surface adapted so that it can be readily covered with colored linings and a concentrated filament electric bulb placed so that an intense and approximately even illumination is given the linings. The details given here are for an arrangement illuminated by a 50 watt type C Mazda bulb. If it is desired to use a larger source of illumination it will be necessary to modify all the specifications given so that the reflecting linings are evenly lighted.

The general lay out will be plain from an examination of Figure 1 which shows a cross section through the centre of the apparatus. A mounting board (MB) about 26 cm. square is fitted with a screw eye (S) along the middle of one edge that the complete arrangement may be hung on the wall at a convenient height. A porcelain base socket (PS) is

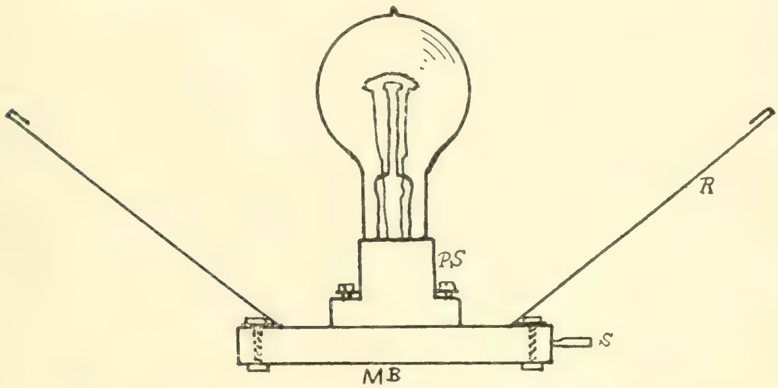


FIG. 1. Semi-diagrammatic cross section through middle of apparatus.

fastened at the centre of one surface of the mounting board and the reflecting surface (R) is mounted so that it is centered around the lamp. Working plans of the reflecting surface are given in Figure 2. This is preferably made of tin and then blackened after assembly. The outer or larger edge of the reflector is turned in to retain the sectors of colored papers which are used to line it and which serve as the stimuli for the larger field of after-images. The radii given are indicated in centimeters. As indicated in drawing 2 a sector of 75° is cut away from the reflector pattern to give the most advantageous angle to the completed surface. The four ears left on the 6 cm. circle are used to fasten the reflector to the mounting board. Care must be exercised in selecting the material out of which the reflecting surface is formed for if it is of a very heavy guage it cannot be formed into the conic frustrum after the retaining groove has been made without rendering the groove useless from "buckling." The dimensions given in Figure 2 may also be used in cutting the colored linings.

Observations may be made in a fully lighted room but the plainest images are obtained in a light proof room after dark adaptation has been accomplished. Here, as in other related observations, a stimulation of a duration so short as to preclude eye movements is desirable and for a simple

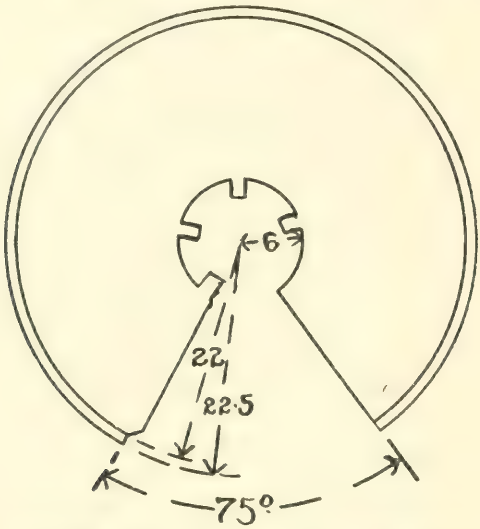


Fig. 2. Working drawing for construction of the reflecting surface.

study of after-images only monocular observation is to be preferred. A group of ten to fifteen can make observations varied and sufficient for class demonstrations or exercises. The most satisfactory and illuminating observations, however, are possible only when the reflecting lining practically covers the entire monocular visual field.

This apparatus gives rise to an almost unparralled series of images. The U-shaped filament of the bulb gives rise to the *first* positive image while the linings give predominantly negative images. Due to the spread of the linings over the entire visual field and the retinal zones a significant series of qualitative changes and temporal as well as spatial differences can be noted even by an inexperienced observer. If it is desired the image of the filament may be practically eliminated by using a type C-4 Mazda in which the bulb is white coated—not frosted.

While this device is especially useful in group demonstrations in introductory courses it also can be put to a very serviceable use in the more advanced and exclusivelly laboratory courses. At the present it is being used for a study of the course of retinal adaptation by means of after-images but its value as a didactic device prompts placing it at the disposal of psychologists in general.

THE MODIFICATION OF THE INTENSITY OF SENSATION BY ATTENTION

BY S. M. NEWHALL

Wesleyan University

I. INTRODUCTION

The general problem of the present report may be stated as follows: what change, if any, in the intensity of sensation is caused by change in attention? The historical part aims to very briefly summarize the available facts and theory of the whole field of the effect of attention on the intensity of sensation. The experimental part has aimed to deal with what may be termed direct effects of attention on sensational intensity. To this end it has seemed possible to factor the attention problem by abstracting from certain of the effects of attention which are physiologically conditioned, such as; the increase in the intensity of sensation which is a usual concomitant of sensory accommodation, and the decrease in intensity which results from sense-organ fatigue. Sensation-changes caused by sense-organ changes, when the latter follow from the functioning of the sense-organ in the attentive process, cannot be said to be due directly to attention.

The experiments were confined to a single sense-field, namely, audition. This is largely accidental and should be regarded as the beginning of a more general experimental attack. The reason why we began the investigation with sound was because binaural hearing seemed to offer exceptionally satisfactory conditions for attention-changes with a minimum of adjustment of the sense-organs, while changes in the position of the binaural phantom seemed to offer unusually clear and unambiguous criteria for changes of relative intensity.

The general problem is of considerable interest and importance. We are still profoundly ignorant of the nervous correlate of attention. Whether it is conditioned by the fundamental neural processes of summation and refractory phase, reinforcement and inhibition, or whether it is exclusively an affair of apperceptive integration we are totally in the dark. The question whether attention conditions any real modification of intensity, and the quantitative question of degree of intensive change, has far-reaching implications in dynamic psychology.[1]

Our immediate experimental problem may be stated as follows: what sort of an effect, if any, do greater or lesser degrees of attention have upon the intensity of an auditory sensation, when the latter corresponds to a stimulus of standard effectiveness as far as the sense-organ is concerned.

II. HISTORY

An historical summary of our problem readily divides into two parts, namely: (*A*) a descriptive account including what have appeared to be significant observed facts, and (*B*) a review of attentional theories which take account of the question of intensive change as due to attention.

A. On the phenomenological side we find a great variety of observation and opinion; and we make a four part division according as the attention has apparently (1) constituted an existential condition for sensation, (2) effected no change, (3) an increase, or (4) a decrease in sensational intensity. These categories seem to include all the possibilities if we let them cover both relative and absolute changes.

1. The evidence from a mass of more or less desultory observation and experiment seems to indicate that some degree of attention is necessary for sensational consciousness. Our experiments tend to prove this. While hundreds of examples of what is meant may be drawn from popular experience; the fact, as it seems to us, has been scientifically described by Fechner,[2] Ulrici,[3] G. E. Müller,[4] and A. Dissard.[5]

[1] Both for the problem and for the general experimental method, I am indebted to Professor Raymond Dodge. I take great pleasure in expressing my thanks for his frequent aid and advice.

[2] 'Elemente der Psychophysik,' Leipzig, 1860, II, 452–3. 'Revision der Hauptpunkte der Psychophysik,' Leipzig, 1882, 270–2.

[3] Cited by Geyser, Joseph: 'Ueber den Einfluss der Aufmerksamkeit auf die Intensität der Empfindung,' Munich, 1897, 14–17.

[4] Cited by Geyser: *Op. cit.*, 17–18.

[5] 'Influence de l'attention sur la perception des sensations,' *Revue Philosophique* XXXIX., 1895, 454–5.

2. The evidence in support of the hypothesis of no change in intensity is relatively insignificant. Fechner,[1] Ulrici,[2] Lipps,[3] and Geyser[4] each concluded that attention has no effect on sensory intensity. Wundt[5] believed that any intensive effect was quite unessential and secondary. Külpe's[6] later experimental results point in the same direction. C. S. Meyers[7] and Felix Arnold[8] are quite undecided concerning the general question as to whether or not there is any intensive change directly attributable to attention.

3. Introspective evidence for an increase in intensity is so varied as to make detailed classification impracticable. A considerable group of writers, however, find a limited increase. G. E. Müller[9] notes intensification of weak sensations and in exceptional cases; while Stumpf[10] finds such intensification to be the rule in audition. Wundt,[11] Nicolai Lange,[12] James,[13] Stout,[14] and Ward[15] also favor a certain limited intensive increase with attention.

There is, further, a group who are inclined to believe that there is some considerable increase in sensory intensity due to attention. Hibben[16] draws a general conclusion of intensive increase from the study of a pathological child. Exner[17] maintains that attention directly increases sensational

[1] 'Elemente d. Psychophysik,' II., 452–3. Revision, 270–2. In 'Sachen der Psychophysik,' Leipzig, 1877, 86.

[2] *Loc. cit.*

[3] 'Grundthatsachen des Seelenlebens,' Bonn, 1883, 134–5. Geyser, *op. cit.*, 25–6.

[4] *Op. cit.*, 57, 103.

[5] 'Grundzüge der Physiologischen Psychologie,' 6 Aufl., III., 315.

[6] 'Der Einfluss der Aufmerksamkeit auf die Empfindungsintensität. Dritter Internationaler Congress f. Psychologie,' Munich, 1897, 180–2.

[7] 'Experimental Psychology,' London, 1909, 319.

[8] 'Attention and Interest,' New York, 1910, 27–8.

[9] Cited by Geyser, *op. cit.*, 18.

[10] 'Tonpyschologie,' Leipzig, 1883, I, 373–5. *Op. cit.*, 1890, II, 290–2, 293–4.

[11] *Op. cit.*, 4 Aufl., II., 274. *Op. cit.*, 6 Aufl., III., 314–5.

[12] 'Beiträge zur Theorie der sinnlichen Aufmerksamkeit,' in Wundt's *Philos. Stud.*, Leipzig, 1888, IV., 391, 421.

[13] 'Principles of Psychology,' New York, 1896, I., 425–6.

[14] 'Analytic Psychology,' New York, 1909, I., 250–1.

[15] 'Psychological Principles,' New York, 1919, 69, 119.

[16] 'Sensory Stimulation by Attention,' *Psychological Review*, II., 372.

[17] 'Entwurf zu einer physiologischen Erklärung der psychischen Erscheinungen,' 1894. Especially 163–171.

intensity, but that this increase by attention is recognizably different from increase due to increase in intensity of stimulation. Ebbinghaus[1] says, "A general increase of the intensity of sensation through attention is . . . thoroughly probable." Titchener[2] favors this view also. While Pillsbury[3] gives the problem considerable attention and states that it is quite unsettled, he seems to favor the positive view.[4] Mach,[5] over thirty years ago, and Bentley[6] about fifteen, came to the experimental conclusion that attention intensifies moderately intense auditory sensation. Bentley, indeed, found the intensification to be relative to the absolute intensity of stimulation,[7] confirming the speculation of Pillsbury.[8]

4. The experimental findings of two psychologists seem to fall in the paradoxical category. While extensive criticism[9] has shown Münsterberg's results[10] to be invalid as far as general significance is concerned; they appear to correctly indicate the decrease in intensity which is often experienced under expectant attention. Tsukahara's results,[11] as interpreted by himself, are opposed to those of Münsterberg; but, as it seems to us, the more correct interpretation indicates a decrease in intensity with attention. Tsukahara's experimentation is open to grave criticism, and therefore must not be taken too seriously.

The available evidence from the data of consciousness is very diverse. While the weight of opinion favors some interdependence of clearness and intensity, the problem cannot be said to be settled.

[1] 'Grundzüge der Psychologie,' I., 586, 588.

[2] 'A Textbook of Psychology,' New York, 1919, 279–280. 'Lectures on the Elementary Psychology of Feeling and Attention,' New York, 1908, 212–220.

[3] 'Attention,' New York, 1908, 2–11.

[4] Titchener, Lectures, 212.

[5] 'Tonpsychologie,' II., 293–4.

[6] Titchener, Lectures, 361–6. Briefly reported in *Psychological Bulletin*, IV., 1907, 212 f.

[7] Titchener, Lectures, 363–4.

[8] *Op. cit.*, 3.

[9] *E.g.*, by Stanley, H. M., 'Attention as Intensifying Sensation,' *Psychological Review*, II., 1895, 55–7. Hamlin, *op. cit.*, 43–66.

[10] 'The Intensifying Effect of Attention,' *Psychological Review*, 1894, I., 39–44.

[11] 'Problem of the Relation of Intensity of Sensation to Attention,' Hiroshima, Japan, 1907.

B. On the theoretical side we also find a variety of views, but the problem of classification is less difficult partly because we shall consider here only attentional theories which have seemed significant in relation to our problem. There are six types of theory to be noted, each one of which makes some attentional factor or aspect (as Pillsbury would say) the causal principle. These explanatory principles may be stated as follows: (1) centrifugal reinforcement of sensory impulses, (2) central psychical reinforcement, (3) prominence in consciousness due to inhibition, (4) prominence due to combined facilitation and inhibition, (5) attention as essentially conditioned by motor factors, and (6) the identity of clearness and consciousness.

1. In attention, centrally originated centrifugal impulses in the afferent pathways summate with the centripetal impulses underlying the sensation, with consequent intensification of the latter. According to Stumpf,[1] this summation process is a common accompaniment of the attentional intensification of weak sensation, and of the raising of subthreshold stimulation to consciousness. In its essentials this theory was also held by G. H. Meyer[2] and G. E. Müller.[3]

2. The reinforcement in this type of theory is purely psychical. The reinforcement may be by congruent memory images, kinæsthesis or other contents of consciousness (Lotze,[4] Wundt,[5] Ulrici[6]); or it may be due to the 'Energie' of the mind itself as a whole (Lipps[7]). The conscious result is in general increased clearness and distinctness.

3. The inhibition theory—both physiological and psychological—increases the prominence of a content by the reduction of the rest of the field. This is the theory of Wundt,[8] Nic. Lange,[2] and Külpe.[10] It involves primarily changes in

[1] *Op. cit.*, I., 374–5.

[2] 'Untersuchungen über die Physiologie der Nervenfaser,' Tübingen, 1843, 250.

[3] From a quotation by Geyser, *op. cit.*, 17–18.

[4] 'Medizinische Psychologie,' Leipzig, 1852, 509. 'Microcosmos,' Leipzig, 1856 (tr. N. Y., 1897), 209.

[5] *Op. cit.*, 4 Aufl., II., 274. 'Outlines of Psychology,' 3d ed., Leipzig, 1907, 238.

[6] Cited by Geyser, *op. cit.*, 15.

[7] *Loc. cit.*

[8] *Op. cit.*, 4 Aufl., II., 274.

[9] *Op. cit.*, 421.

[10] 'Outlines of Psychology' (tr.), 1895, 445.

relative clearness and only concerns intensive modification in an incidental way.

4. Exner[1] combines the theories of facilitation and inhibition and makes the former the more fundamental in attention. This compound theory has a detailed physiological grounding; and its exponent states that attention directly increases sensational intensity. He notes, as stated above, that attentional increase is different from stimulational increase.

5. Ribot[2] is the important exponent of motor theory, and he makes movement the cause and the phenomenon of attention. This theory is open to serious criticism as is also the motor theory of Münsterberg;[3] which is similar to Ribot's though much less thoroughgoing. Motor theory forms the basis for the statement that strain sensations arising in the attention process may be the cause of over- and underestimation of the stimulus attended to.

6. As noted in our descriptive account (*A*), there is a very close relationship between consciousness and attention. Kohn[4] identifies the two. While some degree of attention is always involved in consciousness, fusion of the two concepts would be unfortunate in many respects. Feeling has no clearness.

III. EXPERIMENT

An Investigation of the Effect of Attention on Sound-Intensity

A. A Secondary Objectified Criterion of Change in Tonal Intensity

Our experiments are concerned with the effect of attention on the intensity of noise. The following indirect method was used because it seemed easier and surer than the alternative method of direct comparisons. The observer is presented with a binaural sound and makes a judgment as to change in the location of the phantom when the attention is directed to

[1] *Op. cit.*

[2] 'The Psychology of Attention' (tr.), 3d ed., 1896.

[3] *Op. cit.*

[4] 'Zur Theorie der Aufmerksamkeit,' Halle, 1895. Cited by Pillsbury, *op. cit.*; 292 f.

one component; as compared with its location when the attention is not given a particular direction, or when it is directed to the opposite component.

We provisionally assume that if the phantom shifts with the attention, while the stimuli are in the same phase, of equal frequency, and constant intensity, there is a change in the relative intensity of sensation. This assumption is based on the fact that, aside from asymmetrical auditory fatigue or systemic rhythms, if wave-phase and frequency are constant, the only remaining known determinant of the position of the phantom is the relative intensities of the binaural stimuli.

B. Apparatus

The experiments fall into two main groups differentiated by the procedures followed. The set-up of apparatus is essentially the same in both groups, and it was distributed among three adjacent rooms to be known respectively as *A*, *B*, and *C*. We may perhaps most clearly describe the set-up by tracing the path of the energy-changes which finally produce the auditory stimuli for the observer.[1] These start with a tuning fork and end with telephonic diaphragms. The set-up is diagrammatically represented in Fig. 1.

In room *A* there was an electromagnetically driven Koenig fork (*f*) of 100 V.D. The source of current was three chloride accumulators (*ch*). A D.C. milliammeter (*mi*) and a variable resistance (*r*) made it possible to adjust the intensity of the current. The leads from this circuit passed into room *B*.

In room *B* the leads were attached to the primary (*i*1) terminals of an inductorium situated on the experimenter's[2] table. There were two shunts across the primary terminals, (*a*) a variable resistance and (*b*) a mercury-contact short-circuiting key. The latter normally remained closed, and it was important that its operation cause no disturbing clicks in the receivers of *O*. Across the terminals of the secondary (*i*2) was connected the resistance wire (*w*) of a wheatstone bridge. The sliding contact (*s*) used on the bridge-wire was weighted to maintain an even pressure.

[1] 'Observer' is hereafter designated by *O*.
[2] 'Experimenter' is hereafter designated by *E*.

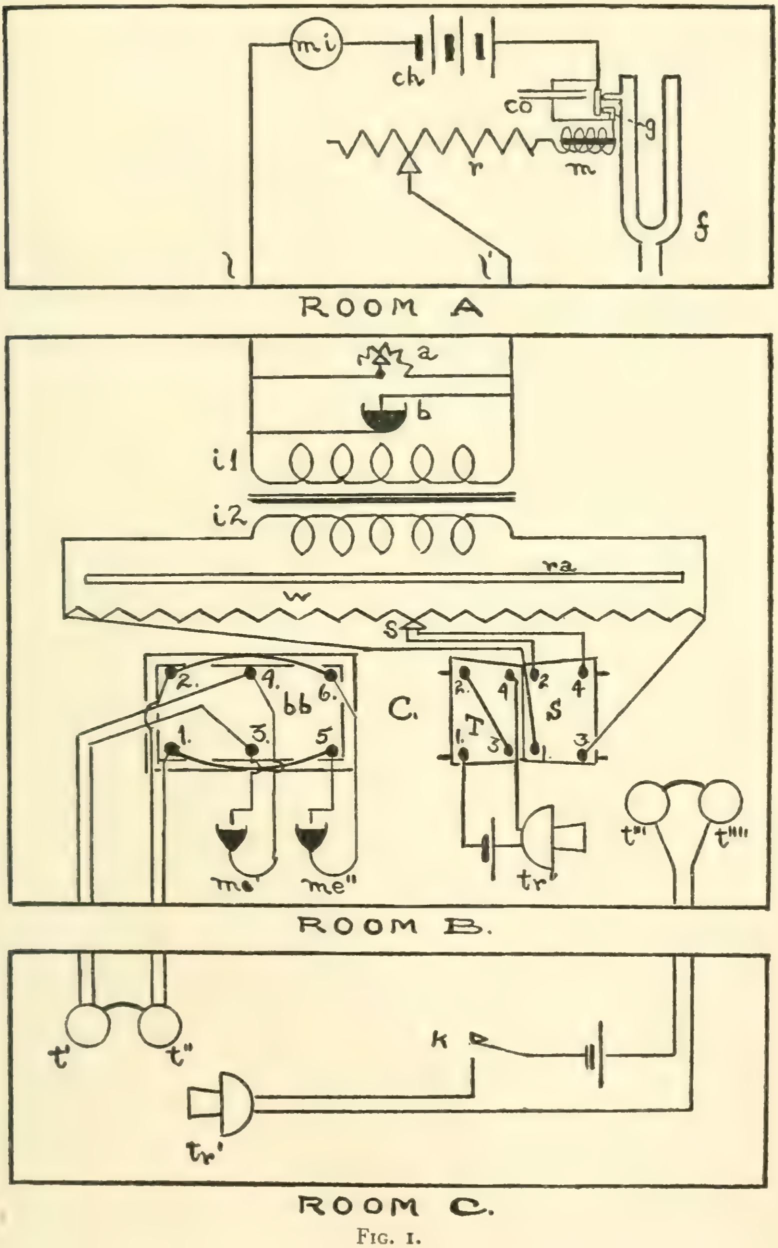

mi
ch
co
g
r
m
f
ROOM A
a
b
i1
i2
ra
w
S
bb
C.
T
S
me'
me"
tr'
ROOM B.
t'
t"
k
tr'
ROOM C.

FIG. 1.

Two flexible wires from the sliding contact and one from either end of the bridge-wire led into a mercury pole-changer (c). They were there so connected with a four-wire line leading to O's receivers in room C that one receiver was connected across the sliding contact and one end of the bridge-wire while the other receiver was connected across the contact and the other end of the bridge-wire. When switch (b) was open, the relative resistances of the receiver circuits could be varied with the sliding contact. It was thus possible to vary the relative amounts of current in the receiver-coils and so the intensity of the auditory stimuli; while the total intensity remained approximately unchanged. Mercury contact shortcircuiting keys (me', me'') made it possible to cut out either of O's receivers.

The time relations of the stimuli were determined within the personal equation of E by operating the shortcircuiting key (b) synchronously with the swinging of a seconds pendulum in E's visual field.

The experimenter was provided with a wireless headset (t''', t'''') for receiving introspection from O, and a mounted telephone transmitter (tr'') for talking to O. By manipulation of the mercury pole-changer (c), E could instantly disconnect O's headset (t', t'') from the bridge-wire and connect it with his transmitter (tr''). This change in connection was accomplished by the use of an eight point contact block. Four of the contacts (S 1234) when inserted in mercury pools of the base block (bb) serve to connect O's headset with the bridge. This is the connection for stimulation. The other four contacts (T 1234) when they are inserted in the base block pools, connect O's headset with E's transmitter. This is the connection for instructions. In regard to the bridge-wire, the connections of O's receivers are interchanged when S 1234 are inserted in the four base block pools to the right hand (bb 3456), as compared to insertion in the four left hand pools (bb 1234).

Six wires pass from room B to room C, two of which lead to each of O's receivers and two to his transmitter (tr'). Considerable care was taken to make the contacts as good as possible in O's receiver circuits. Wherever practicable contacts were soldered or were mercury pools. Since O's transmitter and E's headset were in a separate circuit, O could speak to E whenever desirable. The observer was seated alone in room C which is windowless, and the least noisy room in the laboratory. The transmitter was mounted on a level with his mouth and he closed the telephonic circuit when talking by means of thumb and finger compression on a telegraphic key (*k*). On the white wall opposite O, and about three meters away was a plainly marked fixation point.

C. Procedure and Results

Group I

Procedure.—E presents O with a plainly audible binaural sound of standard intensity. O telephones the apparent location of the phantom to E. E varies the relative intensities of the left and right components and then gives a second stimulation. O telephones the new localization judgment. By continuing this process for one minute or less, E determines the position of the sliding contact which corresponds to localization of the phantom in the approximate median plane. From this contact position and two positions at 5 cm. intervals on either side of it, E constructs a table on his record sheet. For example, if 45 cm. is the median contact position then the other contact positions will be 35, 40, 50, and 55 cm. The five positions determined in this manner were the 'stimulation points' for a complete experimental series or single sitting of O.

A series consists of either three or four parts with 25 'record' stimulations in each part. Five record stimulations are given at each of the five contact settings, following a predetermined chance order. The first part is a 'normal' part, *i.e.*, 25 stimuli of a full second's duration and 4 seconds interval are presented in the chance order. This is with visual fixation in the median plane and no other specific direction of the attention.

In the second part, E presents O with a 1 second left (or right) preliminary stimulus. The preliminary stimulus is followed by the binaural stimulations in chance order as above. *O* is instructed to direct his attention to the preliminary stimulus, and to maintain his attention-set until after the reception of the binaural or record stimulus, the localization of which he judges and immediately phones to *E*. *E* follows the chance order of stimulation and part two of the series is completed.

The interval between the preliminary stimulus and the record stimulus was in 18 parts 2 seconds, in 28 parts 1 second, and in 11 parts 4 seconds. There was thus a total of 1,425 experimental judgments in Group I.

Seventeen of the series were completed with a third part which was normal, *i.e.*, the same as part one; while 20 others contained a second experimental part with a time interval shorter, equal or longer than in the first experimental part. Every series began and closed with a normal part.

In a few series the intensity of the experimental stimulus was increased and in a few decreased. In a few series the *O*'s eyes were allowed to move 'naturally' to the side to which attention was given. The results from these series are insufficient for tabulation but are later discussed.

Seven observers served in the experiments of Group I., *i.e.*, D., N., Di., F., H., S., and W. Introspection was encouraged and with the exception of D. and N. they worked without knowledge. With the above exception, they were all Wesleyan undergraduates taking or having taken two or more psychological courses. Each served for not more than one hour on a given day. Their generous disposition of time and careful work is greatly appreciated.

D. served in 4 series and made 100 record judgments, while N. served in 5 series and made 150 record judgments. Di. and S. each served in 5 series and made respectively 200 and 225 record judgments. F., H., and W. each served in 6 series and made 250 record judgments each. Thus there was a total of 37 series and, as noted above, 1,425 experimental judgments. Since the first and last part of each series was normal, there were 50 normal judgments in each series or a total of 1,850 normal judgments.

Table I

Excess Judgments to the Side of the Attention and the Uniaural Preliminary Stimulus

Observer	Conditions of Stimulation														
	L1		R1		L1, L4		R1, R4		L2, L2		R2, R2		Totals		
	+	−	+	−	+	−	+	−	+	−	+	−	+	−	% Minus Excess
D........	0	11	1	19.5									1	30.5	97
N........	0	16	6.5	4.5	0	12							6.5	32.5	83
Di.......	0	17	1.5	0	0	51	0	22	0	14.5	(*a*)		1.5	104.5	99
F........	0	7	8.5	0	4	6	4	6	3.5	2.5	1	4	21.0	25.5	55
H........	0	4.5	0	12	0	20.5	0	24	0	23	0	15	0	99.0	99
S........	(*b*)		0	5	0	16	0	13	0	22	0	16	0	72.0	72
(*c*) W.......	0.5	0	10	0	33	0	0	52	2	4	22	0	67.5	56.0	45
Sum of totals													97.5	420.0	81%

Table I shows the bias of judgment of position caused by an uniaural preliminary stimulus.

Columns L1 and R1 include the results when the binaural stimulus followed the uniaural Left or Right preliminary stimulus after an interval of 1 second. L2, R2, L4, R4, have analogous meanings.

The plus (+) column gives the excess of judgments in the direction of the uniaural preliminary stimulus. The minus (−) column gives the excess of judgments in the opposite direction. Each figure gives the total excess, plus or minus, judgments resulting from all experiments on any observer under the conditions indicated.

(*a*) Thrown out on introspection which clearly indicated inadequate stimulation.

(*b*) Thrown out on an introspective report of 'tendency to eye-movement' to the side of preliminary stimulus.

(*c*) This *O* is subject to catarrhal trouble. He is the only *O* who made more judgments to the side of the preliminary stimulus than to the opposite side.

The total number of record judgments from which the excess judgments were computed is 1,500 minus the 75 rejected as explained in footnotes, or 1,425 judgments.

Seven per cent. was the excess of the total record judgments to the side of the preliminary stimulus (+), and 30 per cent. was the excess to the opposite side (−). The preponderance of "minus" over "plus" excess judgments is 30.0 divided by (30.0 plus 7.0), or 81 per cent. as shown in Table I.

While the number of excess judgments is, in each case, a small per cent. of the total record judgments, it is significant, with the exception of two cases, that there is a far greater per cent. of excess judgments to the side opposite to that of the preliminary stimulus, than there is to the side of the preliminary stimulus. This relation is strikingly shown in the last column of Table I., where, using in each case the total excess judgments as a basis, the per cent. of judgments to the side opposite to the preliminary stimulus is given.

Results.—The numerical results of Group I. were computed, for each series, in the following way: In part one of a given series, the number of left, center and right judgments were respectively counted. The judgments for the other normal part were counted in the same way. The 'center' judgments of each part were evenly split into the left and right judgments of such part. The average of the left judgments (including one half the center judgments) was then determined; and likewise with the right judgments. The experimental parts were treated in the same way. The sum of the differences of the experimental and normal judgments was then determined. The resulting figures indicate the 'excess' left or right judgments for the experimental parts as compared with the normal parts.

TABLE II

EFFECT OF INCREASING THE INTERVAL BETWEEN THE PRELIMINARY STIMULUS AND THE RECORD STIMULUS

Observer	Conditions of Stimulation							
	(L1, (R1,		L4) R4)		(L2, (R2,		L2) R2)	
	+	−	+	−	+	−	+	−
Di	0	44	0	29	(a)		(a)	
F	4	6	4	6	3.5	4	1	2.5
H	0	25	0	19.5	0	18.5	0	19.5
S	0	18	0	11	0	15.5	0	22.5
Totals	4	93	4	65.5	3.5	38.0	1	44.5
Ex Di	4	49	4	36.5				

The differences between 1 second and 4 second intervals as shown in Table II. indicate that there is a decrease in the number of excess judgments to the side opposite to that of the preliminary stimulus when the inter timulation interval is increased. In the left half of the table the excess judgments following the 4 second interval are in each case, save one, fewer than the excess judgments following the 1 second interval. The results in the right half of the table seem to show that this result is not due to temporal order. For in the right half of the table the interstimulation periods are both 2 seconds and there are a slightly *greater* number of excess judgments on the second record stimulus, whereas in the left of the table there are *less* excess judgments on the second record stimulus.

(*a*) Thrown out because the introspective report indicated that the stimulation was inadequate in series "R2, R2." The total used number of record judgments is consequently 700.

Group II

Procedure.—Effort to eliminate the complicating asymmetrical fatigue effect due to the uniaural stimulus employed in Group I., led to a second group of experiments with the following procedure. Each *O* was given preliminary training in 'hearing-out' the left and right components with the phantom sound in the median plane. This type of hearing-out consists in a voluntary arbitrary emphasis on the given component, with resulting localization of the sound to the corresponding side. The experience of hearing-out is clean-cut and definite. The *O* is quite confident of the fact when he has heard out a given component. The fact of hearing-out seems to indicate a high degree of attention.

In order to eliminate the possible effect of gross eye-movement, the hearing-out in over one half of the experiments of Group II. was done with visual fixation in the median plane. All of the *O*'s found it possible to separate eye-movement and direction of attention.

As in Group I, interstimulation intervals of 1, 2, and 4 seconds were used. Also stimulation at three different intensities was tried, *i.e.*, slightly suprathreshold, plainly audible as in Group I., and moderately intense stimulation. The estimated total amplitudes of diaphragmic vibration are, respectively, 1.52×10^{-7}, 3.12×10^{-7}, and 4.18×10^{-6} cm.[1]

Aside from the above, the procedure throughout was as follows: In any given experimental series, *E* first determines the empirical center and records the five stimulation points as in Group I. Then *O* tries to hear-out of the sound given at the empirical center, the desired component. *O* makes this attempt during a binaural preliminary stimulus of 5 seconds duration. In case *O* does not hear-out in that period, stimulation is discontinued for 5 seconds and then resumed,

[1] These estimates were made from photographing the diaphragmic vibration as communicated to a single taut silk fiber by means of a fine glass rod mounted vertically at the center of the diaphragm. Using the Dodge microscope recorder the shadow of the magnified thread was photographed on moving sensitized paper, and measurement made of the vibrational amplitude in the photograph. The resulting figure was divided by the optical magnification, and corrected for the current flowing in the magnet-coils as compared with the flows under experimental conditions.

and so on, until such time as he does hear-out. It took a practiced O not more than one or two 5 second periods in which to hear-out according to instructions. When O has heard-out the desired component he gives a double key-click, audible to E. E then, after a predetermined interval, presents O with a record stimulus of a full second's duration, following the chance order for Group II. O is instructed to maintain his attention-set throughout the interstimulation interval and until he has judged the record stimulus. After 25 record stimuli have been given the above procedure is repeated while O hears-out the other component. Each series, then, consists of two parts.

Six O's served in the experiments of Group II., *i.e.*, D., N., Di., F., H., and S. D. and N. only served in a few series, and while the results were in the same direction as of the other O's, the number of series was deemed insufficient to warrant tabulation of results. Di., F., H., and S. each served in 13 series and made 650 judgments each.

Results.—The influence of attention to the left and to the right on the relative number of left and right judgments is estimated by Yule's coefficient of association.[1] This is computed as follows: In any given series, the total number of left judgments with attention to the left, and the total number of right judgments with attention to the left are counted. In each part of the series the 'center' judgments are evenly split into the left and right judgments. Then the difference of the products of the judgments which are on the same side as the attention and those which are on the opposite to the attention, are divided by the sum of such products. The result is the coefficient of association (Q).

Since the coefficient is computed for the judgments with the attention to both the left and the right together, it represents the total deviation resultant from attention; and not the bias resulting from attention to the left or right component alone. Under the experimental conditions, we cannot properly separate the left and the right biases; and so compute the coefficients for each separately. It is to be noted that the

[1] 'An Introduction to the Theory of Statistics,' London, 1911, 37–8.

Table III

Judgments Consequent to Hearing-Out One Component of a Binaural Tone

A. Eye-Movement to the Same Side

	Observer Di.					Observer F.				
Heard out—	Left		Right			Left		Right		
Interval	+	−	+	−	Q	+	−	+	−	Q
1″........	15.5(a)	9.5	16.0	9.0	.49	19.5	5.5	21.0	4.0	.90
1″........	14.5	10.5	17.5	7.5	.53	22.0	3.0	23.0	2.0	.98
4″........	16.0	9.0	14.0	11.0	.39	17.0	8.0	17.5	7.5	.66
4″........	15.0	10.0	12.0	13.0	.13	13.0	12.0	22.5	2.5	.84
2″........	13.0	12.0	19.0	6.0	.55	18.5	6.5	16.5	8.5	.69
2″........	11.0	14.0	19.0	6.0	.46	14.0	11.0	16.5	8.5	.42
Totals.....	85.0	65.0	97.5	52.5	.42	104.0	46.0	117.0	33.0	.78

B. Central Visual Fixation

	Observer Di.					Observer F.				
Heard out—	Left		Right			Left		Right		
Interval	+	−	+	−	Q	+	−	+	−	Q
2″........	11.5	13.5	17.0	8.0	.29	14.0	11.0	14.5	10.5	.28
1″........	12.0	13.0	14.5	10.5	.12	14.5	10.5	16.0	9.0	.42
4″........	16.5	8.5	14.0	11.0	.42	14.5	10.5	15.0	10.0	.35
2″ (b).....	13.0	12.0	19.0	6.0	.55	15.5	9.5	14.0	11.0	.35
2″........	14.0	11.0	18.0	7.0	.53	17.5	7.5	11.0	14.0	.29
2″........	14.0	11.0	15.0	10.0	.31	12.5	12.5	15.5	9.5	.24
2″........	13.5	11.5	14.0	11.0	.20	18.0	7.0	14.0	11.0	.53
Totals.....	94.5	80.5	111.5	63.5	.34	106.5	68.5	100.0	75.0	.34

(*a*) This table gives the total plus and minus judgments in each part-series of 25, instead of excess judgments as in Tables I. and II. For example, observer Di. had 2 series with an interval of 1 second between the preliminary and record stimuli. Out of 25 cases where he heard-out to the left he made 15.5 left judgments and 9.5 right judgments, in the first series. Under the same circumstances, in the second series, 14.5 were judged to the lef. and 10.5 to the right.

(*b*) The stimulus for the first and third of these last four series was weak for each *O*, and for the second and fourth, it was moderately intense. (*Cf.* next page.)

Table III

Judgments Consequent to Hearing-Out One Component of a Binaural Tone

A. Eye-Movement to the Same Side

Heard out—	Observer H.					Observer S.				
	Left		Right			Left		Right		
Interval	+	−	+	−	*Q*	+	−	+	−	*Q*
1″........	16.5	8.5	15.0	10.0	.49	16.5	8.5	13.0	12.0	.36
1″........	11.5	13.5	10.0	15.0	−.28	14.0	11.0	16.5	8.5	.42
4″........	18.0	7.0	11.5	13.5	.37	12.5	1 .5	18.5	6.5	.48
4″........	15.5	9.5	13.0	12.0	.[illegible]8	14.0	11.0	16.0	9.0	.39
2″........	14.0	11.0	11.0	14.0	.00	15.0	10.0	17.5	7.5	.56
2″........	13.0	12.0	12.5	12.5	.40	14.0	11.0	21.0	4.0	.44
Totals.....	88.5	61.5	73.0	77.0	.15	86.0	64.0	102.5	47.5	.49

B. Central Visual Fixation

Heard out—	Observer H.					Observer S.				
	Left		Right			Left		Right		
Interval	+	−	+	−	*Q*	+	−	+	−	*Q*
2″........	15.5	9.5	14.0	11.0	.35	12.5	12.5	16.5	8.5	.32
1″........	13.0	12.0	12.5	12.5	.40	14.0	11.0	17.0	8.0	.48
4″........	17.5	7.5	9.5	15.5	.18	15.0	10.0	15.0	10.0	.38
2″ (*a*).....	14.0	11.0	13.0	12.0	.16	15.0	10.0	17.0	8.0	.52
2″ (*b*).....	13.0	12.0	18.0	7.0	.47	17.0	8.0	12.5	12.5	.36
2″........	13.0	12.0	17.0	8.0	.39	15.5	9.5	18.5	6.5	.65
2″........	9.5	15.5	18.0	7.0	.22	16.0	9.0	13.0	12.0	.32
Totals.....	95.5	79.5	102.0	73.0	.31	105.0	70.0	109.5	65.5	.40

Comparison of the individual coefficients of Tables III., *A* and III., *B* would seem to show that there is no significant correlation between the M.V. and eye-movement to the side heard-out, for the limited number of cases here presented.

The first two 'intensity series' (*a*) and (*b*) are comparable inasmuch as they were both made on the same day, one following the other. This was true in the case of each *O*. The same was true of the second pair of intensity series. The estimated intensity of the total binaural stimulus used in the first series of both pairs is 0.000 000 152 cm. and in the second series 0.000 004 178 cm.

Six out of the eight comparisons, *i.e.*, two for each *O*, show a higher coefficient for the weaker stimulus. This would seem to show that the attention insofar as the attention-set consequent to hearing-out is concerned, is more effective in the case of the weak than the stronger binaural sound used.

Table IV

The Group Totals of Record Judgments for Group II, and the Corresponding Coefficients of Association between Attention to One Side and Judgments to the Same Side

A. Eye-Movement to the Side of Attention

Heard out Left			Right		Judg't Totals	Q
Observer	+	−	+	−		
Di	85.0	65.0	97.5	52.5	300	.42
F	104.0	46.0	117.0	33.0	300	.78
H	88.5	61.5	73.0	77.0	300	.15
S	86.0	64.0	102.5	47.5	300	.49
Totals	363.5	236.5	390.0	210.0	1200	.48

B. Central Visual Fixation

Heard out Left			Right		Judg't Totals	Q
Observer	+	−	+	−		
Di	94.5	80.5	111.5	63.5	350	.34
F	106.5	68.5	100.0	75.0	350	.34
H	95.5	79.5	102.0	73.0	350	.31
S	105.0	70.0	109.5	65.5	350	.40
Totals	401.5	298.5	423.0	277.0	1400	.34
Sums of Totals	765.0	535.0	813.0	487.0	2600	.41

Inspection of column Q reveals several interesting relations. The coefficient of any O as found in IV., A, is appreciably different from that in IV., B. For three of the O's the coefficient in IV., B, is considerably less than in IV., A, while for the remaining O it is greater. It would seem that in cases where we have eye-movement to the side of the attention, there is an increased tendency to judge the phantom as on the same side.

While it is perhaps not legitimate to make inferences from the fact, the group totals show an interesting relation. Whereas in IV., A, they are widely different, in IV., B, they are within .09 of each other. The M.V.'s are respectively .18 and .03, or as 6 to 1. The decreased deviation may be partially a result of the elimination of individual differences due to eye-movement.

coefficient is greater or less than it would be if a wider or narrower range of relative intensities in the binaural components were used. Consequently our figures have no absolute or exact quantitative value. In a few supplementary experiments, however, it was found that if the binaural phantom as physically conditioned, was at one limit of the particular range of relative intensities which we used, say to the left, it was barely possible to hear-out the right component even though it was physically conditioned by the weaker stimulus. This was roughly true for all the *O*'s tested, and serves to give our figures a little more than mere relative value.

However, the most significant thing about the coefficient of association is that in every series, with two exceptions for observer H., it indicates some degree of shift to the side to which the attention is directed.

IV. CONCLUSION

The important difference in the conditions of our two groups of experiments was, that in Group I. the preliminary stimulus was uniaural and served as an aid to the direction of the attention; whereas in Group II. the preliminary stimulus was binaural and the attention was voluntarily and arbitrarily directed to a given component. We may interpret the results of the groups separately and then state their combined indications.

1. The results of Group I. clearly indicate a biased sensory fatigue-effect resulting from the employment of an uniaural preliminary stimulus (Table I.).

2. As far as its influence on the position of the phantom is concerned this fatigue effect is a very delicate and variable phenomenon.[1] The experiments with different interstimulation intervals suggest a recovery from fatigue, or an increase in intensity of sensation, when the interval between the uniaural and the record stimulus is changed from 1 to 4 seconds.

3. The results of the few experiments on the effect of

[1] These results are in harmony with those of Flugel, J. C., 'On Local Fatigue in the Auditory System,' *British Journal of Psychology*, XI, 1920, 133–4.

different intensities of stimulation indicate much greater fatigue at the weaker intensity.[1]

4. The results of the few experiments with eye-movement to the side of uniaural stimulation were markedly different from the rest which employed central visual fixation; and seem to indicate that eye-movement tends to counteract the uniaural fatigue effect, as far as localization judgments are concerned.

1. In Group II. both the hearing-out and the attention-set which follows it cause a change in the position of the phantom.

2. The attention-set appears to be a variable as regards duration; but several indicators suggest that it persists for some time. For instance, in over half of the tested cases, the empirical center was found to be farther away from the side to which the attention was directed, at the middle than at the beginning of the series. Also, the *O*'s frequently reported that it was more difficult to hear-out the opposite component in the last half of the series. Two *O*'s felt very sure that their judgment was influenced by the preceding hearing-out. One remarked that there seemed to be a 'hangover' which increased the difficulty of hearing-out the opposite side. All *O*'s reported that after hearing-out one side for a time the process becomes easier and the phantom becomes located more or less continuously nearer to the side of the attention (*i.e.*, the empirical center is farther to the other side). One *O* reported that sometimes, well along in a series, the phantom is already at the to be heard-out side at the commencement of the hearing-out periods. The writer had this experience also.

3. The attention-set may be shifted, apparently, contrary to instructions by a relatively small change in the external conditions. From the introspective reports of the *O*'s and his own record sheet, it was clear to the experimenter that after several cases in which the binaural components were so different that the phantom was judged to be on the side opposite to that heard-out, it was more difficult than usual

[1] *Cf.* Bowditch in Verworn, Max, 'Irritability,' 1913, 164.

to hear-out according to instructions in the subsequent preliminary period.

4. The results of the experiments with different intensities of stimulus seem to show that the attention-set has a greater effect on weak than on strong intensities of sound. Observers occasionally reported increase in total intensity after the experimental series had commenced, which would seem to indicate sensory accommodation or adaptation.

5. One of the most striking introspective facts is the experiencing of various rapid 'shifts' and 'spiral' movements of the phantom. Each *O* experienced something of this sort, though comparatively infrequently. From its character and the sequence of the changes, it seems extremely improbable that the cause was instrumental. "When hearing-out the right, . . . the stimulus will appear, at times, to shift from center to left, and then back to right where it remains." "At times the 5 second stimulus (the preliminary binaural stimulus) shifts from center to left and back to center, sometimes only once, sometimes two or three times." "The shift is a clean-cut and amazing phenomenon, distinctly objective; there is a true illusion of motion; it moved clear across the center to the other side." The above are introspections from three different observers. Sometimes, in hearing-out, the phantom is reported as shifting part way over only to 'slip' back again.

Combining our interpretations, we find five conditions which effect the judgment of position of the binaural phantom, when wave-phase and frequency remain constant, *i.e.*, absolute intensity, relative intensity, sensory fatigue, eye-movement, and attention-set.

1. In both groups of experiments stimulation near the absolute threshold of intensity seems to favor the shift of the phantom.

2. With regard to the direction of the attention, the shift is in opposite directions in the respective groups.

3. The effect of eye-movement is similar in both cases, *i.e.*, an increased tendency to judge to the side of the attention.

4. The shifts and spiral effects which were noted much the more frequently in Group II., according to the introspection, would seem to be due to changes in attention.

Especially interesting is the fact that at the weaker intensity of stimulation we found greater shifts of the phantom both as due to sensory fatigue and to attention-set. The results of our experiments confirm the findings of Stumpf,[1] Nic. Lange,[2] G. E. Müller,[3] and others, *i.e.*, that attention has most effect on weak sensations. Sometime ago Bowditch[4] showed that recovery of irritability after weak stimulation is slower than after strong stimulation. If fatigue and inhibition are both dependent on the refractory phase as Verworn maintains, these two apparently opposed phenomena may both be expressions of the same fundamental neural fact.

We are under obligations to Professor Titchener for pointing out that our experiments do not preclude a shift of the binaural phantom due directly to changes in clearness. But whether the changes in our record experiments are due to clearness or not the fact that it is possible to hear-out one side of the binaural stimulus to the exclusion of the other certainly shows that as a consequence of attention, intensity of audible tone may be reduced to zero.

While our experiments, then, cannot be said to settle the question of whether attention increases the intensity of the component to which it is directed or decreases the intensity of the component from which it is abstracted, the evidence indicates a decreased intensity of the component from which the attention is abstracted. We have no introspective evidence that the heard-out component becomes more intense. As far as these experiments go, change in the relative intensities of the two components would appear to be due to an absolute decrement rather than an absolute increment.

In regard to the general problem, we cannot safely reason from one sensory field to another by analogy. What holds for auditory sensation, in regard to change in intensity, may be expected to hold for other mental facts of the same class. But this is a matter of fact to be settled only by experimentation in the various sense fields.

[1] *Op. cit.*

[2] *Op. cit.*

[3] *Op. cit.*

[4] Cited by Verworn, *loc. cit.*

AN IMPROVEMENT IN VOICE KEYS

BY KNIGHT DUNLAP

The Johns Hopkins University

The voice keys I designed earlier for use in the timing of association reactions were simplified and improved from the Römer key, and have always given some trouble in the adjustment of the contact. The difficulty has arisen from the method of leading in the current to the swinging lever. Sufficiently good conduction cannot be obtained through the bearing of the lever, and I therefore employed a fine wire connection directly from the lever to a binding post. Such a wire connection can be made to work, but if the wire becomes bent (as it easily may, and frequently does, through removal of a cloth cover, or by accidental touches), it changes the pressure of the lever against the diaphragm, breaking the circuit permanently or decreasing the sensitivity of the key. With improvements in the chronoscope, the voice key has remained the only weak point of my Association Timing Apparatus.

The final model of the voice key, illustrated in the cut, removes the last difficulty, offering a key in which the contact is operated by gravity alone, and in which the maximum of sensitivity with maximum of reliability is therefore obtained.

The key consists, as in earlier models, of a brass ring (*S-S*) supported by a brass block (*P*) on a base (*B*) of hard rubber or bakelite. Within the ring is supported a diaphragm (*D*) of thin aluminum, carrying at its center a gold contact plate mounted on a mica disc (*F*). The contact-lever (*L*), with its gold-pointed head (*H*), is pivoted in a bearing (*G*) attached to the top of the ring (*S*): thus differing from the early models, in which the lever was pivoted at the bottom. From the top of the lever extends a bent wire (*W*) dipping into a mercury-cup (*C*).

slightly in the mercury, no trouble is encountered at that point. The key is highly sensitive and satisfactory, and may be considered a final model.

The nose-plate (N) attached to the top of the ring (S-S) is useful in the cases of untrained reactors, who tend often to speak *around* the key, when the experimenter is opposite the reactor. By instructing the reactor to touch the tip of the nose to the plate (N) at the command 'ready,' this source of difficulty is easily obviated. With reactors accustomed to the apparatus, the nose-plate is not needed.

The contact-plate (*F*) on the diaphragm is connected to a binding post (*J*) by means of a thin wire (not shown). The mercury-cup (*C*) is connected with a similar binding post (not shown) through the ring (*S-S*). In this way, the current is

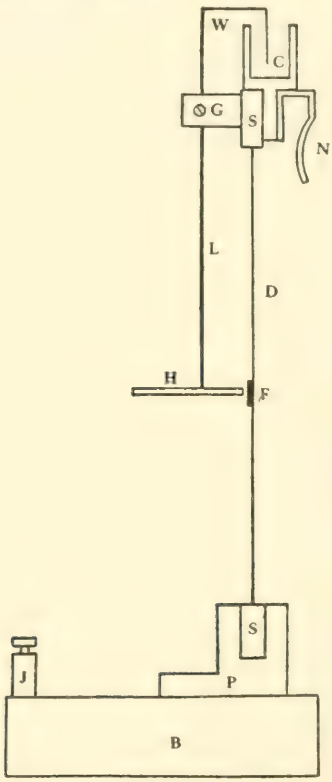


Fig. 1. Cross section of key. *D*, aluminum diaphragm, suspended in ring *B* with supporting block *P*, on base *B*. *L–H*, contact-lever, pivoted at *G*, and with bent extension wire *W* dipping in mercury-cup *C*. *N*, nose-plate. *F*, contact-plate on diaphragm.

led in to the lever, and yet the lever is perfectly free to act under the force of gravity and the vibration of the diaphragm; and by adjusting the inclination of the whole key (by means of a leveling screw in the supporting tripod, not shown), the required pressure of the gold tip of the lever-head against the gold contact-plate can be secured. Since the wire (*W*) does not lift out of the mercury cup (*C*), but merely moves

Journal of
Experimental Psychology

VOL. IV, No. 4. AUGUST, 1921

THE LATENT TIME OF COMPENSATORY EYE-MOVEMENTS

BY RAYMOND DODGE

Wesleyan University

The latent time of a reaction no longer takes the pre-eminent place in psycho-physical description that it once assumed. But while we have in large measure outgrown our earlier enthusiasm for reaction-time measurements, it still remains true that no other single characteristic of a human reaction carries so many implications with respect to the probable origin and the classification of the act within the system of known psycho-physical processes as its latency. Probably also no other factor constitutes such a delicate objective control of the relative constancy of human experimental conditions. We must still measure the effects of experimental biases like drugs, habituation, fatigue, and attention largely in terms of their effects on the reaction time. It is still on the latency of a reaction that we must depend for our estimate of the degree of neural complication that is involved in its elaboration, and for our estimate of the neural level of its origin in normal life.

The relatively long latency of the saccadic eye-movements, for example, as well as the long latency of pursuit movements stamps these types of eye-movements with the character of complex reactive processes, removes them from the class of reflex responses, and places their probable origin within the superior neural systematizations.

There is no satisfactory data concerning the latency of the compensatory eye-movements. This is not explained merely by the fact that hitherto no suitable recording device has been available. Even with our delicate mirror-recorder fully developed, the technical difficulties in recording the latent time of reaction to bodily rotation still remained almost prohibitive. These difficulties were due chiefly to the stimulus inertia. It is extremely difficult to bring the body mass into sudden motion. Moreover, it is still impossible to state the moment of acceleration when the rotation stimulus reaches super-threshold proportions. The obvious expedient under such circumstances is to produce a maximum acceleration as suddenly as possible. Unfortunately this proved not to be so desirable as it seemed. Most devices to produce sudden acceleration of large masses involve both noise and jar. In eye-reaction experiments, however, clicks, snaps, and all excessive vibration must be carefully avoided since they tend to produce lid reflexes that directly or indirectly obscure the onset of eye-movements.

Notwithstanding an intermittant experimental attack lasting many months, it is doubtful if our least objectionable onset of bodily rotation ever gave us a sharply defined moment of stimulation at all comparable to the standard reaction techniques. It may confidently be asserted that the latency of eye-reaction compensatory to bodily rotation has not yet been measured with the accuracy that characterizes our measurements of the knee-jerk, the lid-reflex, speech-reactions, or even the latency of the other eye-reaction types. Moreover, it is doubtful if a stimulation process that involves movement of the body mass can ever reach such refinement. There is, therefore, ground for supposing that the true latent time of compensatory eye-movements is less than the recorded time between the beginning of body rotation and the beginning of eye-movement. Such measurements should be regarded as approximations only. They give us the general order of the latency, not its exact measurement.

In view of the difficulties involved in starting the body-mass sufficiently quickly and silently to serve as a satisfactory stimulus moment, the idea arose of rotating the head alone. This reduces the mass moved. But we have not yet found a technique for completely relaxing the neck muscles. There is a constant tendency to actively interfere in the process, resisting or facilitating the passive movement according to some unknown condition. The best reflex records of this sort that we have succeeded in obtaining were made by a method that we shall describe in the second part of this paper. They are best described after the technique. It is sufficient here to say that they show a reflex time in the lower part of the distribution curve of reflex measurements.

Since it is quite likely that no one who subsequently attacks the problem of recording the reflex times of compensatory eye-movements will be entirely satisfied with the method we finally adopted, an outline of promising techniques which proved unsatisfactory may prove as useful to the next experimenter as a description of the technique that we used. In that spirit we shall briefly describe the development of the rotation technique.

In recording the compensatory eye-movements of rotation it is apparently essential that the recording camera, the source of illumination, and all accessory apparatus rotate with the subject. To accommodate this paraphernalia we substituted for the customary revolving chair a small turn-table. It was mounted on a heavy tripod base which was screwed solidly to the floor.

The bearings consisted of a central shaft about 10 cm. long, accurately cupped at the bottom, resting on a small hemispherical point of hardened steel. This point carried the main gravitational thrust. Side thrusts due to more or less inaccurate balance of the load, were taken by a ring of polished steel balls running in the customary turned grooves with screw adjustment. For studying the threshold of rotation it later seemed desirable to further diminish vibration by increasing the side thrust supports. Wire guys were consequently carried from the four corners of the turn-table to a hollow ball-bearing which was supported about three meters from the floor by struts from the wall. These three bearings constituted a remarkably smooth-running and silent system of supports, quite rigid and fairly free from vibration when the turn-table was suddenly started and stopped.

The turn-table consisted of a horizontal platform about 100 cm. by 60 cm., and about 20 cm. from the floor. On this turn-table are built as rigid and as light as practicable, a bench seat which runs lengthwise of the platform, and an apparatus table. Changes in the height of the seat are made on the primitive but effective plan of adding or withdrawing pads of paper. The apparatus table rises from one end of the platform to a height of 75 cm. It carries not only the recording camera but also the head rest, the lighting system with its resistances, and such other apparatus as the experiment required.

The recording camera consisted of the improved Dodge-Cline camera. A narrow slit in front of the sensitized paper provided for the linear orientation of simultaneous parts of the record. Time ordinates were introduced into the record by two devices. The simplest means was the periodic introduction of resistance into the lighting circuit. This produced periodic fainter spots in all the record lines, which appear in a straight line running across the record and indicate simultaneous moments of the record as well as the time. This method was used only to indicate two second intervals. The period was controlled by the seconds pendulum of an accurately running clock, acting through a relay. For more frequent time ordinates a fork-driven synchronous motor was used. This carried a perforated disk which was interposed between the camera slit and a secondary source of illumination. By this device dark line ordinates of any desired frequency could be written across the record simultaneously with the record curves.

Each reaction record comprises three record lines in addition to the time ordinates. These may be designated as follows: (1) The rotation record, a line which indicated the onset of rotation and the acceleration of the turn-table. (2) The head-line, to show how much if any of the eye-line displacement may be due to head-movement. (3) The eye-line, the main record line which must show the moment of eye-reaction and the subsequent course of the eye-movement.

The Rotation Record

Head and eye-lines were provided for by means of the mirror-recorder which was described in the previous paper. The development of an adequate rotation record proved to be a real technical problem that delayed our reaction measurements for weeks. When it was finally solved the method seemed obvious enough. Meantime we had tried out in fact or in imagination a considerable variety of unsatisfactory devices.

We assumed that the ideal rotation record would be made by a concave mirror, like the mirror of the eye-movement recorder. Ideally this should be situated on the vertical axis of the turn-table and attached immovably to some fixed point of the environment. No support could be arranged from the top of the turn-table without interfering with the wire supports and the wires to the illumination system. This was equally and even more obviously true of supports from the sides. No support could come from the cone bearing at the bottom without destroying its delicacy. These difficulties were finally avoided by a shaft rising from an excentric axis beside the tripod support, and terminating in a cone bearing at the head rest. The shaft carried a piece of concave mirror of slightly less focal length than the recording mirrors of the eye and head, since it stood a little in front of them though in the same horizontal plane.

To yield a record of rotation it was necessary for this shaft to remain motionless while the turn-table revolved around it. This was provided for by a friction belt from a pulley on the tripod base to a pulley of the same diameter on the excentric shaft. We made this belt of electrician's tape. It is ideal for such purposes, extremely flexible, and consequently without internal friction. It can be tightly adjusted to prevent lost motion, has absolutely no slip, and operates silently. This device solved the rotation record problem perfectly. In later experiments where it became necessary to drive certain mirror and shutter devices, a chain gear drive was substituted for the tape. With more work to do the tape had a tendency to stretch.

The Onset of Rotation

The onset of rotation of the turn-table as we have already noted was a matter of special solicitude, vital to the accurate measurement of reaction latency. Since the stimulus onset should be both rapid and silent we decided at the beginning against an electric motor drive for reaction experiments, and in favor of a spring. Heavy rubber bands proved satisfactory. They operate positively, evenly, and silently. They can be increased in number and tension to meet any requirements of force.

The real difficulty was the release of the turn-table. We were unable to devise any trigger release that was both silent and rapid acting. We tried several. Magnetic release seemed plausible. In operation, however, it could not be divorced from click and jar. The former seemed to be inherent in the demagnetization of the big core. The most satisfactory release that we achieved, *i.e.*, the quickest with the least jar and noise, was dependent on the skill of the operator rather than on a mechanical device. It worked as follows: Until the moment of release, the operator held the table still, against the tension of the springs by pulling an offset arm which was firmly attached to the table, against a fixed stop, with a padded hook. At a predetermined moment the operator shoved the hook away from him with the utmost possible suddenness. This movement did not actually rotate the table. It would have been too irregular. It merely permitted the stretched spring to exert its force against the inertia of the turn-table. It was impossible to prevent a certain amount of vibration, but with care the onset could be made noiseless, relatively vibrationless, and surprisingly uniform under constant tension of the spring. The operator's task was to get the padded hook out of the way faster than the table started to move, without hitting anything.

It was not easy to reset the springs to the same tension for different directions of rotation, and on different experimental days. We did not try for exactitude in this respect. This is a fault of the present series, though a minor one. The actual acceleration rate of each record is shown directly in the record itself by the slant of the rotation line. In view of the method of release it did not seem advisable to try for greater refinement.

To reduce refractoriness and sensory fatigue in these experiments, the turn-table was not rotated indefinitely but was caught by the rubber springs after about 20° of rotation, gently decelerated, reversed, and finally stopped by hand.

The resulting records show that the latency of ocular compensation to rotation differs considerably from individual to individual, and apparently in the same individual according to attention, though this latter is not satisfactorily explored. The instructions were to attend to some fixed spot in space (usually some part of the room or outside landscape that seemed directly in front of the subject), to keep looking toward that spot when rotated, but without opening the eyes.

Figures 1, 2, 3, and 4, Plate I., reproduce four records of compensatory reaction. In records 1 and 2 the subject was rotated counter-clockwise, in 3 and 4 clockwise. These records should be interpreted as follows: They were taken on a falling record plate. The horizontal time ordinates progress upward, 0.01″ apart. That is, the record reads from below up. The rotation line *S* shows as a double curve that leaves the vertical slowly and passes off the record in a single uninterrupted sweep. Since the recording beam actually stands still while the turn-table revolves, the direction of rotation is always opposite to the direction of the rotation line. The eye-line *R* shows a sudden break from five to six time lines after the rotation begins, in the opposite direction to that of the rotation line.

The subsequent more or less irregular course of the eye record need not now concern us. The first significant movement is compensatory. Since both the rotation line and the eye-movement line reverse the actual direction of movement, their mutual relation is correctly represented by the records. The head-line, on the contrary, and the head-movement factor in the eye-line are recorded in their true direction. With the beginning of rotation the head line shows a slight inertia lag. The eye-reaction curve emerges sharply from

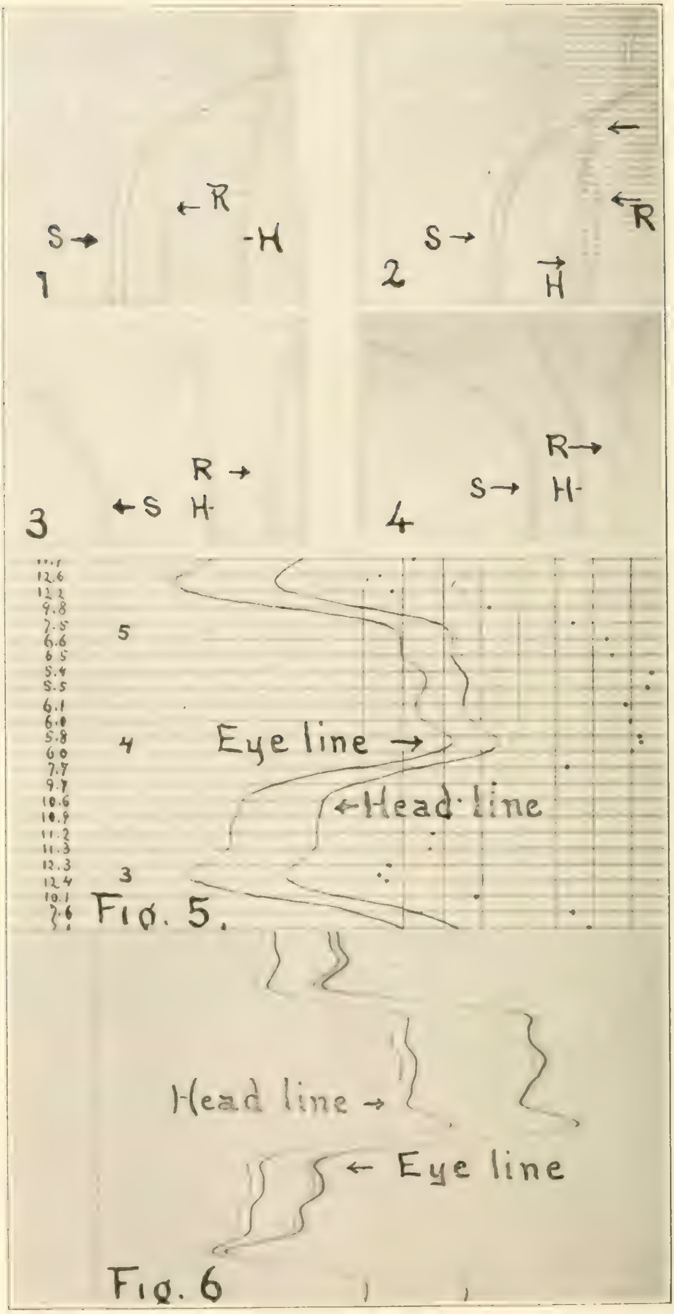


Plate I.

Figs. 1–4 are records of reflex compensatory eye-movements, time ordinates 0.01″ apart.

Figs. 5–6 are records of coördinate compensatory eye-movements, time ordinates 0.05″ apart.

this phase, clearly defining the moment of reaction. It is never so easy to determine the exact moment of the beginning of rotation for reasons that we have already fully considered.

In record number 2 a notable break in both head and eye-curves may be seen about 0.09″ after the beginning of the compensatory eye-reaction. This is obviously a record of head-movement that interrupts the simple inertia drift. It is further obvious that the corresponding phase of the eye-movement curve is of greater amplitude than the head curve. The eye-movement in this case cannot be a simple translation of the eye by the head. Some sort of eye-muscle action has changed the relative position of head and eye-mirrors. This is a curious phenomenon and will further engage our attention in the second part of this paper. If we regard this eye action as some kind of compensatory reaction to a voluntary head-movement we must note a conspicuous difference between it and the reflex compensatory eye-movement which occurred 0.09″ earlier. Its latency is either zero or very close to it. We shall be able to demonstrate this peculiarity even more clearly in records that are planned to isolate it experimentally.

At the end of this head-movement episode the reflex compensation continues as though it had not been interrupted. About 0.10″ later the compensation is interrupted by an eye-movement of the first type, the rapid or saccadic type. After this the compensation curve flattens out as the oscillatory movement of the turn-table approaches its maximum amplitude and slows down, preparatory to beginning the return swing. These parts of the record are cut away to save space. The accuracy with which such changes in the direction of rotation can be followed by compensatory eye-movements together with some remarkable adaptation phenomena will be discussed in a later paper.

Returning to the latency of the compensatory movements, I have plotted all available measurements in Fig. 7. With our present technique only one record can be taken on a plate. This is not conducive to quantity production. In addition, the stimulus difficulties which we have already discussed, make the multiplication of records in the effort to establish norms, of doubtful expediency.

The measurements which we have plotted in Fig. 7 were taken from six subjects, all undergraduate students of psychology except myself. I have more legible records than any other subject. They are represented in the distribution figure by crosses. Unfortunately many of the records of the other subjects proved illegible on account of excessive head movements, lid reactions, or accidental displacement of the recording light above or below the camera slit. All legible records are represented in the figure. The units are hundredths of a second.

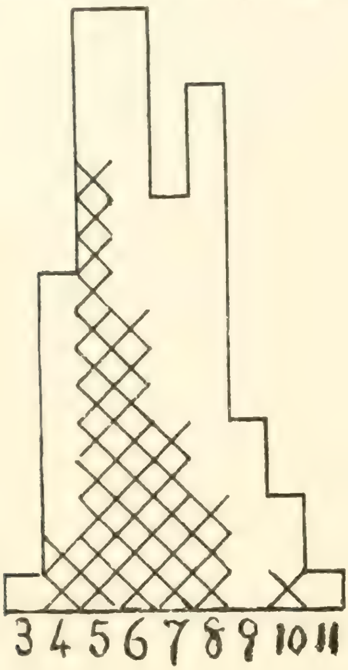


FIG. 7 is the distribution area of all available measurements of the latent time of reflex compensatory eye-movements of six subjects. The units are hundredths of a second.

There is little chance of misinterpreting these data. The mass of the measurements lie between 5 and 8 hundredths of a second. This seems to include some at present unaccountable individual differences. My own mode is 0.05″. These values are clearly the latent times of sub-cortical reflexes. The latency of saccadic eye-reactions is from three to four times as long.

Without considering the pathological evidence, if we had no other data to depend on than the present records, these differences in the reaction latencies of the two phases of the compensatory process, the true compensatory and the saccadic, would indicate an essential difference of origin. This indication is borne out by evidence which will be given in a later paper, that the saccadic phase of the compensation is the more accessible to voluntary control and training.

PART II

The Latency of the Coördinate Compensatory Eye-Movements

In record number 2, Plate I., we found indications of a different kind of compensatory eye-movement, compensatory to voluntary head-movement rather than to passive rotation. This type of eye-movement had one remarkable characteristic, distinguishing it not only from saccadic but also from the reflex compensatory eye-movements; its latent time was either zero or very close to it.

Since the eyes may remain open during these 'coördinate' compensatory eye-movements, they offer fewer technical difficulties to photographic registration than the reflex type. In my 'Five Types of Eye-movements'[1] I described a method for photographing them, together with the resulting kinetograms. Supplementary experiments by the Lamanski method which were described in the same paper proved that their continuous coördination with the head-movements to which they are compensatory, was remarkably accurate. Unfortunately the records by the technique which was then available, were not clear enough to be reproduced. So it is perhaps not surprising that notwithstanding the evidence, the existence of these remarkable eye-movements as a special type has not been generally admitted.

Our new mirror-recorder permitted us to make some extraordinarily clear records of this third type of eye-movements. The chief difficulty in photographing them is the requirement that the head be free to move. This requirement reverses the usual desideratum of a motionless head, and precludes the use of the customary head-holding devices. Nevertheless, if the recording images are to fall on the slit, some means of controlling the position of the head must be used. For the earlier records I used an optical control. For these later ones with the mirror-recorder I used a simple forehead rest. In small angular movements of the head, the loose skin of the forehead seems to furnish an ideal, almost

[1] *American Jr. of Physiol.*, 1903, 8, 307 ff.

frictionless bearing. In movements of larger amplitude we used a broad head-band around the forehead, which worked over a friction pulley behind the head.

Theory of Coördinate Compensatory Eye-Movement Records

The chief difficulty in interpreting records of compensatory eye-movements that occur during voluntary head-movement, lies in factoring out of the curves the part that is due to head-movement and the part that is due only to eye-movement. Whenever the head rotates, the eye-globes are necessarily carried with it. If an eye were fixed rigidly in the head, the head-mirror and the eye-mirror would write parallel curves, since both would traverse equivalent arcs on concentric circles without changing their relative positions. For example, if both head and eye started at their primary positions, with the visual fixation point at an infinite distance, rotation of the head 10° to the right would carry a rigidly fixed eyeball 10° to the right as part of the head, and the line of regard would move to the right with the eye. If instead of being fixed in the head the eye were to maintain its fixation point at an infinite distance, in compensation for a head movement of 10° to the right, it must rotate 10° to the left.

On first thought this might seem to indicate that the record line of an adequately compensating eye would remain straight, whatever the angle of head-movement. This would be approximately true if the eye and the eye-mirror both rotated in the same direction with the same angular velocity. With our method of recording, the actual record is quite different. When the apex of the cornea rotates to the left, compensating for a rotation of the head to the right, a superposed mirror resting on the lid over the apex will be rotated to the right. It will regularly rotate in a direction opposite to the direction of eye-movement. That is to say, rotation of the head in any direction will cause the recording mirror of a compensating eye to rotate still farther in the same direction. A compensating eye, therefore, will write a record that is neither straight nor parallel to the record of head movement, but a curve of greater amplitude than the head-movement curve, with homologous changes of direction.

The true compensatory component of the eye-movement curves may be found by subducting from the amplitude of the eye-movement curve the amplitude of the head-movement curve. Whenever the head- and eye-movement curves are parallel, both recording mirrors must have maintained the same relative positions. That will always indicate that the eye remained fixed in its socket. Whenever the curves cease to run parallel, there must have supervened some independent rotation of the eye-globe. If the eye-curve is of greater amplitude than the head-curve, but in the same direction, the eye-movement is compensatory. The moment when parallel curves diverge or converge is the moment of eye-reaction. The interval that elapses between the beginning of a head-movement curve and the divergence or convergence of the record-lines is the latent time of the compensatory eye-movement.

Typical compensatory eye-movement records are reproduced in Figs. 5 and 6, Plate I. In Fig. 5 the long oblique lines were too faint to reproduce and have been retouched. Measurements were taken before the record had been retouched. Fig. 6 is exactly as taken. The two records are from different subjects. One degree of head-movement displaces the head-line approximately one centimeter.

Comparison of the head- and eye-lines in Figs. 5 and 6 shows the relationship that we have just described. The eye-line follows the head-line point for point, but with increased amplitude. At no point is the head-line vertically straight. Nowhere do head- and eye-lines run parallel. This means that while the head is in continual motion, more or less rapid, the eye-compensations are equally continuous. In the attempt to express this compensation quantitatively I measured the separation of the head- and eye-lines on each time ordinate of the record from which Fig. 5 is taken. These measurements appear in the figure as a column of digits at the left. Their values are plotted in a series of dots on the right of the record. These values show a remarkably close correlation between position of the head-line and the separation of the head- and eye-lines.

The time-interval between a change in the direction of the head curve and the divergence or convergence of the two lines would give the latent time of the compensation. Simple inspection of the curves shows that this latent time must be very short. At the end of movement 3, for example, between the preceding time-line and the apex of the curve is approximately one third of the space between time-lines. Since the time-lines in Fig. 5 are 0.05″ apart that would represent a time-interval of approximately 16 thousandths of a second. Yet in that time-interval the separation of the head- and eye-lines has measurably changed. This indicates a latency of considerably less than 0.016″. How much less cannot be estimated so easily. The facts are even more conspicuous at the end of movement 4, Fig. 5.

In some respects Fig. 6 is the best record of the phenomenon which this technique has yielded. While it unfortunately lacks legible time-lines, the speed of the record is approximately the same as that of Fig. 5. Aside from this lack the record is unusually clear. The close juxtaposition of head- and eye-lines together with the wide amplitude of the eye-movement factor favor direct interpretation from simple inspection. In this record the relative obliquity of the two lines both at the beginning and at the end of each movement is conspicuously different, apparently from the very onset of the head-movement.

While inspection of the records is sufficiently convincing to make it clear that the probable latency of the coördinate compensatory eye-movements is of a different and much smaller order of magnitude than the eye-reflexes of vestibular origin, it seemed unlikely that these differences in obliquity could be read with sufficient accuracy to prove whether the latency was ever really zero. The effort to get still more accurate records led to a substantial improvement in the technique which will presently be described.

The Accuracy of Coördinate Compensation

In addition to indicating a very low latency these records confirm our previous observation that the coördinate compensatory eye-movements differ from the reflex compensatory movements and from the ordinary pursuit movements in their smoothness and freedom from corrective eye-movements of the saccadic type (type 1). Such a fine coördination as is here disclosed is without a known parallel in our nervous system. If the records were not so unmistakable it would be incredible. We can understand a nervous mechanism by which the onset of a voluntary head-movement will coördinately produce an eye-movement of appropriate amplitude in the opposite direction. There are numerous familiar examples of such compensation. The coördinate compensatory movements of the limbs which are involved in maintaining equilibrium are conspicuously adequate in spite of vigorous kicking or striking. The peculiar difficulty in the present case arises from the evidence for proportional, adequate compensation for all phases of the head-movement, including the complex antagonistic process involved in stopping the rotating head-mass in a series of damped sine waves. These irregular stops can be seen in both Figs. 5 and 6.

The end of a rapid movement of finger; arm, or head is never a simple stop. The member does not reach an extreme position and stay there as one might expect it would if the contraction of a single agonistic muscle reached some arbitrary limit and held rigidly. There is always a more or less complicated rebound. Part of this rebound is doubtless a matter of momentum acting on elastic muscle tissue. There is evidence that the stop is further complicated by the interaction of antagonistic muscle groups on the basis of reciprocal innervation.[1-2] We have called attention to similar phenomena at the end of saccadic eye-movements.[3]

Whatever the cause, the fact is that no two head-movements end alike. Some show one, some two, and some three

[1] Isserlin, 'Ueber den Ablauf einfacher willkuerlicher Bewegungen,' Lpz., 1910, Engelmann.

[2] Sherrington, 'The Integrative Action of the Nervous System,' N. Y., 1906, Scribner's Sons.

[3] Dodge, 'An Experimental Study of Visual Fixation,' *Mon. Suppl. Psy. Rev.*, No. 35.

oscillations of varying extent before the head comes to rest. A nervous mechanism that is capable of producing coördinate compensation for these irregularities of stopping is quite unknown to us. That such a mechanism exists our records leave no possible doubt. If the technique for recording them could be simplified they would seem to be most favorable points of attack in the study of the effect of drugs that may be supposed to disturb coördination.

A Skeleton Recording Camera

Our efforts to develop a technique for more adequately recording the latency of coördinate compensatory eye-movements were based on two desiderata. First, it seemed desirable to write the curves at higher speeds of the recording paper. That was merely a matter of more brilliant illumination. Secondly it seemed worth while to try to eliminate more or less completely, the head-movement factor from the eye-movement curves. Apparently this could be accomplished only by moving the record plate with the head. It proved a beautiful technical problem, and is not yet entirely solved.

The picturesque ideal would be to attach the whole recording apparatus to the beam of light from the head-mirror so that it would move exactly as the head-record moved. While this is of course impracticable, it may well function as a mental standard. It might seem at first thought that a close approximation to this ideal would be obtained by attaching the recording apparatus rigidly to the head. Such a solution would be very faulty. In the first place since the angular displacement of the recording beam from a rotating mirror is twice the angular displacement of the mirror, the record surface may not move simply with the head by any rigid attachment, but must move approximately twice as fast. Second, if the ordinary recording apparatus were attached to the head the increased inertia and momentum would render the habituated coördinations quite inadequate. One could learn little of normal motor coördinations by such an instrument. The latter difficulty would be exaggerated by the disadvantageous leverage that must result from having the camera at sufficient distance from the head to get a reasonable amplitude of eye-movement record. Both these difficulties, however, can be met more or less satisfactorily by a suitable modification of the apparatus. The mass of the recording apparatus can be reduced to a minimum by rebuilding it in skeleton form. A system of levers can be introduced between the head and the camera to magnify the head-movement to any desired degree.

The successive refinements of apparatus as one difficulty after another was more or less satisfactorily overcome in a gradual approach towards our picturesque ideal are not germain to the present paper. In its present form the skeleton recording camera which best met the requirements may be described as a system of light levers, one end of which carries the sensitive paper. The other end terminates in a mouth-piece that is held rigidly between the teeth of the subject.

The plate-holder is reduced to a cardboard back of just sufficient size to hold the record straight by the use of paper-clips. This holder is hung on the end of a pendulum about two and one half meters long. The pendulum swings the sensitized paper behind a slit in the middle of a large opaque wooden screen 30 cm. in front of the subject.

The post from which the pendulum is suspended, rests on our oil-resistance, moving plate-holder whose various speeds provide for the successive exposure of the paper behind the slit at any desired rate. Even with a long pendulum, the angular changes in the relative position of record sheet and slit as the pendulum oscillates would vitiate the time relations of different parts of the record. That is to say, as the pendulum oscillates the time ordinates would be written at different angles. To prevent this distortion and to keep the record perpendicular to the slit at all positions of the pendulum, the pendulum-rod was doubled. The axes of the two rods were separated horizontally about 20 cm. They were held apart the same distance at the bottom by the skeleton plate-holder. The top of the record thus makes the base of a parallelogram, and will always be parallel to the horizontal line of the axes whatever the position of the pendulum. This insures rectangular orientation of all points on the record. There remains only one distortion of the record due to the movement of the pendulum. The top of the record will be higher at either end of the swing than it is at the middle. If the pendulum is long enough this may be reduced to a negligible error. Given a pendulum that is $2\frac{1}{2}$ meters long, a total excursion of 4 cm. represents a maximum angle of $30'$ with the vertical. Calculating the corresponding change in elevation by the formula $x = (1 - \cos 30')$ 250 cm. gives 0.01 cm. Even if the maximum excursion of 4 cm. took place within the time represented by two time-ordinates, *i.e.*, 0.01″, such a small change in elevation would have no measurable effect on the temporal sequence of the various record curves.

The oscillations of the pendulum and consequently of the sensitized paper were synchronized with the head by means of a system of light wooden levers that increased the angular velocity of the record by two. A light rod was held fast in the mouth of the subject by firmly setting the teeth into the yielding wood. At right angles to this mouth-rod a horizontal rod extended to the left and articulated with the middle of a short vertical lever from whose top another rod connected with the bottom of the pendulum. All connecting rods were made of very thin strips of redwood. Their bearings were thin steel V-shaped knife-edges which were held in contact by elastic rubber bands. In order to minimize distortion of the straight lines of the record as the pendulum fell behind the slit, the horizontal connecting rods were made as long as practicable, about 150 cm. The system was so light and frictionless that subjectively it seemed to have no disturbing effect on the head movements. The records showed, on the contrary, that its momentum did in fact somewhat distort the otherwise delicate compensatory eye-movements.

Figs. 8 and 9, Plate II. show two typical records. The time-ordinates are 0.01″ apart. The head-lines are almost

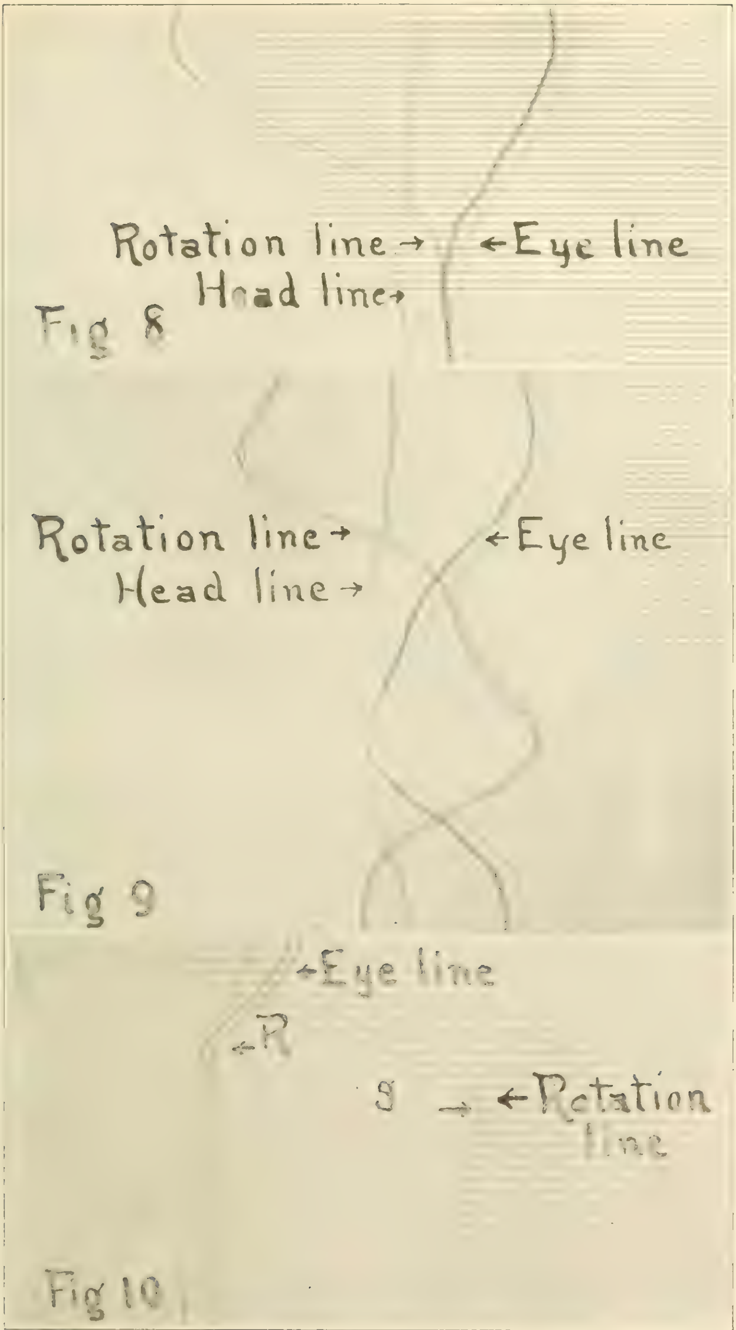


PLATE II.

FIG. 8–9 are records of coördinate compensatory eye-movements, time ordinates 0.01″ apart.

FIG. 10 is a record of the latency of a reflex compensatory eye-movement, time ordinates 0.01″ apart.

straight, fulfilling the intent of the lever system. The extent of head-movement is shown by a third line, the 'rotation line' which is traced by a motionless beam of light as the pendulum oscillates with the head. No latency within the limits of measurement is found at the beginning of head-movement in these two records. In no record was there more than 0.01″ discrepancy. In some the latency was apparently very close to 0.01″. In no case did the eyes start to move first. They commonly stopped moving before the head stopped, and regularly failed to correspond as closely with the course of the head-movement as they appear to do when the head is not loaded with the mass of the recorder.

The skeleton recorder may be said to prove our previous contention that the coördinate compensatory eye-movements may have a latency of zero. It does not fill the need of a convenient instrument for studying the accuracy of compensation under varying psycho-physical conditions. We still believe that optical means may finally be made to displace the mechanical lever system.

Reflex Latency with the Skeleton Camera

The success of the skeleton recorder in measuring the latency of coördinate eye-movement suggested the possibility of measuring the latency of the reflex compensatory eye-movements due to vestibular stimulation, by forcibly rotating the head instead of the body mass. It was hoped in this way to produce a more sharply defined moment of stimulation. The plan worked only moderately well. It seems difficult or impossible to rotate the head of a subject by pulling on the mouth lever of the lever-system without setting the neck muscles in action. Sometimes they operate to resist the pull, sometimes to facilitate it. It proved very difficult to keep them relaxed. Whenever the neck muscles act they complicate the record with coördinate compensatory eye-movements. Furthermore, it proved difficult to hold the head steady before moving it, and to move it suddenly without jar. The best head rotation stimulus that we produced was by the aid of an elastic rubber device, analogous to that which we found most satisfactory in moving the turn-table.

Two rubber bands were attached to the mouth-piece lever and stretched in opposite directions so as to hold it in an elastic equilibrium. No simple disturbance of that equilibrium by a push or pull proved at all satisfactory. Releasing one of the springs by a trigger device produced a decided snap and consequent lid reflex. Some excellent records were obtained by displacing the elastic equilibrium to one side or the other relatively slowly and then suddenly withdrawing the pressure. The elastic that was most stretched then became the motive force for rotating the head in the direction of the elastic equilibrium. It operated suddenly, evenly, and silently.

A record which was obtained by this procedure is reproduced in Fig. 10, Plate II. We believe that this record represents the best available reaction technique for measuring the latency of the eye-reactions to vestibular stimulation. In this record the latency is exactly 0.05″. It falls in the lower part of the distribution of latency measurements, but in the mode of the measurements for R.D. (See Figure 7.)

As a result of these records we may make a critical comparison of the two types of compensatory eye-movement with the pursuit eye-movements.

1. With respect to origin: The stimulus to ocular pursuit-movement is visual and consists of real or apparent motion of the object of regard. The stimulus for reflex compensatory movements is rotation. Its receptor field centers in the vestibule. Whether other origins of the perception of rotation can produce compensatory eye-movements has never been proved, though it is a question of great theoretical and practical importance. The stimulus for coördinate compensatory movements is central. In our present ignorance of the mechanism it may be hypothetically pictured as a collateral impulse of the same origin as the impulse that moves the head.

2. With respect to latency: The pursuit movements regularly begin after a natural reaction time of the order of 0.2″. It may be considerably reduced by practice and a limitation of the range of possible directions of motion. The

reflex compensatory movements have a latency of the order of 0.05″. The coördinate compensatory movements may have a latency of zero, but faults of coördination up to 0.01″ appear in our records.

3. With respect to onset: The pursuit movement begins with a series of inaccurate approximations to adequate pursuit, interrupted by saccadic movements either in the same or opposite direction, usually the former. Only after practice does the pursuit movement become an adequate, smooth pursuit. On the basis of 'short-lived motor habits,' when adequate pursuit of a pendulum is once originated, it tends to persist within the limits of fatigue, with only small and infrequent corrective movements. Persistent pursuit in one direction is rhythmically interrupted by saccadic movements that bring the eye back to approximately its primary position. This produces a kind of nystagmoid-motion that cannot be objectively distinguished in the records from the reflex nystagmus of rotation. Reflex compensation shows less hesitation in its onset. Since there is apparently no subjective criterion of adequacy when the eyes are closed, the greatest variety of compensation speeds may result from the same rotation stimulus. Whether or not compensation can be eliminated by practice, whether it is ever adequate are still open questions. It is regularly interrupted under normal circumstances, like the pursuit movements, by saccadic returns to approximately the primary position of the eyes, giving rise to the phenomenon of reflex ocular nystagmus as we may call it to distinguish it from the nystagmus of pursuit. Unlike both the other eye-movements, coördinate compensation movements are astonishingly smooth. They are never interrupted, under normal circumstances, by corrective movements except as the point of fixation is voluntarily changed. For moderate amplitudes of head-movement and moderate velocities, coördinate compensatory eye-movements are amazingly accurate.

4. Interaction of the various types: When the eyes are open during passive rotation of the head, pursuit must coöperate with reflex compensation, for clear vision of the

environment. We have no knowledge of how the coöperation is effected. Similar coöperation will doubtless obtain between pursuit and coördinate compensation. Voluntary rotation of the head which conditions coördinate movement of the eyes must also act as stimulus for reflex compensation. In our records there is no indication of the latter type interplaying with the former. One might expect a reaction phenomenon at the end of the normal reflex latent time, but no such disturbance of the smooth coördinate compensation has been found.

In some of these combinations the factors may on occasion conflict. There are various indications that such conflicting tendencies are a source of discomfort. As to which tendencies win out in such conflicts, and what the characteristics of the resulting eye-movements would be, we are profoundly ignorant except in one case, namely, when the stimulus of decelerated rotation conflicts with the visual tendency to fixation, as in after-nystagmus.

Many of the characteristics of the reflex compensatory eye-movements are also still in doubt. We know neither how adequate the compensation may be, nor how the adequacy may be affected by training. In the case of coördinate compensation there must be differences due to changes in accommodation for distance, as well as for various head-loads. The mechanism of these adaptations is an unexplored field.

AN EXPERIMENTAL STUDY OF ASSOCIATIVE INHIBITION[1]

BY LINUS W. KLINE,

Skidmore School of Arts, Saratoga Springs, N. Y.

Introduction

The several laws of inhibition of the higher mental processes, referred to as associative, retroactive, terminal, etc., have been derived and demonstrated, in part, by the use of nonsense syllables. The laws formulated from such studies, useful as they may be for pure science, are made to appear much simpler and less involved than they doubtless are in meaningful associations.

The law of associative inhibition, first expressed by Müller and Schumann, (8) asserts that the bond of connection between any two cognitive elements tends to inhibit the formation of a similar bond between either and a third element, or "If *a* is already connected with *b*, then it is difficult to connect it with *k*; *b* gets in the way."

A number of studies on the interrelation of growing habits more or less opposed to each other tend to show that they become mutually inhibitive, only under certain conditions. Bair (1) has shown that the interference of two opposing habits formed during the acquisition of skillful movements grow less the more automatic the two responses become, and that "finally, after both sets of reactions were practiced to perfection, the interference disappeared." Culler (4) in his study on 'Interference and Adaptability' points out that "when two opposing associations, each of which excludes the other, are alternately practiced with one, four, or eight repetitions of each association before the other is resumed, the opposing associations have an interference effect upon each other, in

[1] This study was made possible through a fellowship by courtesy granted to the writer by The Johns Hopkins University for the academic year 1919 and 1920, and for which I hereby make grateful acknowledgment.

all" (subjects). The interference effect grows less and less, while the practice effect becomes greater so that either of them can be called up independently without the appearance of the other." Kline and Owens (7) in their study, 'Learning Skillful Movements,' found that many errors of the second series were the correct distributions of the first. Such errors were most likely to occur when the speed was forced beyond the power of the delivering habit. Brown, as quoted by Swift (13), found in an experiment of sorting cards that changing the methods of sorting increased errors and checked speed, but 'that as the work proceeded practice in one order helped in learning the other.'

These experiments are cited here because they emphasize two points: (1) That the amount of inhibition mutually exerted by two opposing movements depends upon the degree of automatism possessed by each, (2) that two opposing movements when highly automatized, may facilitate, rather than inhibit each other. The operation of such laws in more or less opposing skillful movements raises the question if some such condition may not obtain in associative mechanisms involving only central factors, so-called.

May not different degrees of strength in the association bonds of central mechanisms produce decided variations in their inhibiting and facilitating relations? May not the associations involving meaningful material and functioning in daily life conform to other laws than those obtaining in the association bonds for nonsense material?

The present investigation was made to determine what variations, if any, exist in the law when meaningful associations formed in daily life are involved.

Material

Meaningful material conformable to the Müller-Schumann formula must have two terms whose associative bonds are formed by daily experience and the second pair of terms between which associative bonds are *to be* formed must have one term that occurs in the first pair and the other, similar in nature to the second of the first pair, must be void of associations with either term of the first pair. A State of the Union and its capital, *e.g.*, Tennessee-Nashville, matches fairly well the first pair of terms *a* associated with *b*. But if it be attempted to associate a false capital with the state, the condition of *a* associated with *k* is set up, *e.g.*, Tennessee-Sacramento. Two groups of 15 states, each so chosen as to represent all sections of the Union were formed and given false capitals.

The results derived from the use of states and false capitals called for the use of meaningful material derived from other sources to serve both as a check and as a means for following up certain clues bearing on what appeared to be new aspects of associative inhibitions. Two lines of material were selected for this purpose, numbers and false sums—$6 + 4 = 11$—and literary works and false authors. For a check on the association between states and false capitals, entirely unknown counties and county seats were used; for a check on numbers and false sums, the addition of two letters and their sums as $d + h = 17$ were used, while unknown works of literature coupled with their true authors were compared with the literary works and their false authors. The Johns Hopkins chronoscope was used to measure the association time of individual subjects and the stop watch to measure time in the group tests.

Procedure

The material, when arranged in convenient series for learning, was presented to the subjects in groups and as individuals. The method of work with the groups is reported first.

A. Procedure with Groups

College students, men and women, were subjects. They were first asked to write the names of states with which they were familiar (familiar being defined as knowing something of the history, industry and geography of a state, and, in addition, it may mean personal associations of one sort or another). It was found that 58 percent of the capitals known by the men were those of 'familiar' states, and 54 percent of the capitals known by the women were of 'familiar' states.

So, being familiar with a state was by no means a guarantee that its capital would be known, although the condition is some better than fifty-fifty. The 'current knowledge' of the capitals of the fifteen states used in the investigation was ascertained by having each student write from memory the capitals in blank spaces opposite their respective states. (By 'current knowledge' is meant knowing the correct capital of a state on the instant and without previous warning as to the question.) To insure that the student associated the true capital with its state *at least once*, both state and capital were recorded by the student from dictation. See Table I.

Table I

Current Knowledge of States and Capitals—College Students

Familiar States				True Capitals Known: 15 States			
Men		Women		Men		Women	
Frequencies	States	Frequencies	States	Frequencies	Capitals	Frequencies	Capitals
1	1	1	0	1	0	4	0
1	2	1	1	5	1	11	1
6	3	4	2	8	2	15	2
5	4	10	3	9	3	14	3
3	5	13	4	6	4	20	4
8	6	9	5	3	6	12	5
6	7	9	6	2	7	6	6
3	8	16	7	4	8	9	7
2	9	10	8	3	9	3	8
1	10	5	9	2	10	3	9
1	11	6	10	2	11	1	10
2	12	3	11	1	14	2	11
5	13	3	12				
1	16	2	14				
		3	15				
		2	16				
		2	16				
		2	17				
		1	19				
Totals.....45	314	100	714	46	219	100	405
Mean..................	6.97		7.14		4.54		4.05
Mode..................	6		7		3		4
Mdn....................	5.8		6.16				4.3

Read first line thus: One man was familiar with only one state, one woman with no states; one man knew no state capital, four women knew no state capital, etc.

After having secured the data for 'familiar' states and for 'current knowledge,' the learning series (Ser. 1, Group *a*) composed of fifteen states and false capitals, arranged in parallel columns, was presented. The subjects were instructed to associate each false capital with the state on the same line and to learn the two so as to be able to recall the capital if required. The subjects were not told the time allowed for learning the series—fifteen minutes—but that they would be expected to recall them in some form. Subsequent work showed ten minutes to be a better duration for learning a false series.

1. *Immediate Recall.*—All responses of whatever sort throughout the work began and stopped at prearranged signals. Immediate recall was made by writing names of false capitals opposite their respective states in a different order as given in Series I., Group *b*. Time for recall 1′ and 30″.

Series I., Group *a* Arranged for Learning		Series I., Group *b* Arranged for Immediate Recall	
State	Capital	State	Capital
Florida	Columbus	Arkansas	————
Delaware	Providence	Michigan	————
Michigan	Boston	Wyoming	————
Washington	Austin	Iowa	————
Wyoming	Augusta	Texas	————
Iowa	Nashville	Connecticut	————
Vermont	Columbia	Maine	————
Texas	Des Moines	Alabama	————
Connecticut	Trenton	Vermont	————
Pennsylvania	Atlanta	Illinois	————
Colorado	Albany	Delaware	————
Maine	Cheyenne	Colorado	————
Arkansas	Bismarck	Pennsylvania	————
Illinois	Raleigh	Florida	————
Alabama	Phoenix	Washington	————

2. *Recognition of False Capital.*—After an interval of 48 hours, the same list of states arranged in a new order was again presented for the purpose of recognizing the false capitals. Each false capital was placed opposite its state, along with the names of three or four other large cities. Sample lines are here given.

Series I., Group *c*

Strike out with a cross mark thus, ×, the false capital that you have just learned. Make no other marks.

State	Capital
Washington	San Diego, Toledo, Austin, Wheeling, Kansas City
Pennsylvania	Atlanta, Richmond, Pittsburgh, Omaha, Lowell
Florida	Miami, Columbus, Springfield, Dayton, Leavenworth

Sixty seconds were allowed for recognition. The fact of recognition was shown by the subject striking out false capitals.

3. *Delayed Recall.*—The college women made a delayed recall of false capitals after an interval of seven days, recalling 43 percent; the college men made a recall after twenty-seven days, recalling 18 percent. This long interval was from necessity rather than choice.

The names of fifteen counties and their respective county seats were presented, learned, recalled, and recognized under conditions similar to those given above. Table II. shows the number of subjects in each series.

Table II

Showing Results of Immediate Recall of False State Capitals and of True County Seats (15 in Each Series)

15 in Each Series	Men		Women	
	False Capitals	County Seats	False Capitals	County Seats
No. of Subjects	46	41	100	99
Total Number Recalled	464	466	830	12.73
Mean	10.87	11.36	8.30	12.85
Median	10	12	8	14
P.E. of Median	1.230	0.901	0.772	0.613
Percentages	67.20	75.70	55.30	85.00

This should be read: In a series of 15 false capitals 46 subjects recalled 464, or 67.2 percent.

Series I., Group *d*, of Counties

Learning Series

County	County Seat	County	County Seat	County	County Seat
Quitman	Belain	Pike	Magnolia	Alcorn	Corinth
Leflore	Greenwood	Hinds	Raymond	Benton	Ashland
Barbour	Clayton	Pearl	Poplarville	Wilcox	Camden
Copiah	Hazelhurst	Tate	Senatobia	Homes	Lexington
Tippah	Ripley	Sharkey	Rolling-Fork	Webster	Walthall

The results of these three forms of response to false capitals: immediate recall, recognition, and delayed recall, are shown in Table III. They indicate that states whose true capitals were known offered apparently less inhibition than states having capitals unknown to the subjects, that is, unknown at the time under the conditions of the test. Or, in terms of inhibition, a state whose capital is known offers less resistance to being associated with a false capital than a state whose capital is unknown. Table III. shows that the ratio of false unrecalled capitals to known state capitals is, for each of the three forms of response, less than the ratio of false unrecalled to unknown state capitals.

Table III

Results of Inhibiting Effects of Known and Unknown State Capitals on Learning False State Capitals (15 Capitals in a Series)

College Men

No. of Students	Nature of Response	True Capital	False Capital	Ratio of False Unrecalled to Known and Unknown True Capitals
46	Immediate	Known 218	Unrecalled 52	0.2385
46	Recall	Unknown 472	Unrecalled 165	0.3491
46......	Recognition	Known 218	Unrecognized 10	0.0458
46		Unknown 472	Unrecognized 56	0.1144
38.......	Delayed Recall	Known 180	Unrecalled 128	0.7111
38	Delayed Recall	Unknown 350	Unrecalled 339	0.9685

College Women

No. of Students	Nature of Response	True Capital	False Capital	Ratio of False Unrecalled to Known and Unknown True Capitals
100	Immediate	Known 405	Unrecalled 160	0.3950
100 ..	Recall	Unknown 1095	Unrecalled 502	0.4584
93	Recognition	Known 335	Unrecognized 10	0.2985
93......		Unknown 967	Unrecognized 44	0.4550
93......	Delayed Recall	Known 372	Unrecalled 192	0.5161
93	Delayed Recall	Unknown 1023	Unrecalled 599	0.5855

Further investigation was suggested by these results, using association material in which the strength of the bond of connection could be roughly determined *in advance* of the experiment. For such material the field of literature was chosen. Three lists of literary works and their authors, 'well-known,' 'less known,' and 'unknown,' were prepared and varied until satisfactory. The different degrees of familiarity were determined by submitting the lists, composed of twenty or more literary works, individually, to a number of educated persons. They rarely failed to give an author of the supposed well-known list. Several authors of the less known list could not be recalled, or were known to only a few, others were doubtful. Two of the individuals had never connected Edwin Markham with The Man With the Hoe, or Browning with The Ring and the Book. The third list played its part perfectly as 'unknown literature.' The two lists of 15 literary works and their false authors, chosen from this preliminary survey, are here submitted.

Well-known Literature	False Authors	Less Known Literature	False Authors
The Raven	Darwin	Modern Painters	Dante
Les Miserables	Shakespeare	The Ring and the Book	Irving
David Copperfield	Longfellow	Sleepy Hollow	Stowe
Gulliver's Travels	Dickens	She Stoops to Conquer	Ruskin
Faust	Homer	Tam O'Shanter	Huxley
Origin of Species	Stevenson	Vanity Fair	Emerson
Paradise Lost	George Eliot	The Man with the Hoe	Kipling
Pilgrim's Progress	Hawthorne	Vision of Sir Launfal	Whitman
Treasure Island	Milton	Uncle Tom's Cabin	Holmes
Ancient Mariner	Bunyan	The Scarlet Letter	Markham
The Lady of the Lake	Whittier	The Jungle Book	Prescott
Hamlet	Poe	Rasselas	Lowell
Hiawatha	Coleridge	The Inferno	Thackeray
The Iliad	Goethe	The Skylark	Browning
Snow Bound	Tennyson	The Excursion	Sheridan

(In using this literature with large groups, it soon appeared that the degree of familiarity, or difficulty, should have been determined in a preliminary test on a much larger number of subjects not to be used in the experiment proper, of course, and the results arranged in array for purposes of selection. But the present arrangement, while not the best, served the end of this investigation.)

The literature and authors were used, in the main, like the states and capitals. First: The degree of familiarity with the fifteen authors of well-known literature was determined. Second: Fifteen false authors were associated with the well-known literature. Here the learning time was ten minutes. Third: Immediate recall, time 1′ 45″. Fourth: Delayed recall was made 48 hours later, time 1′ 45″. The 'less known' literature was presented and treated similarly. The recognition test (Otis method) was not used with the literature series. The 'unknown literature' was treated as the other two, with the exception of the familiarity test. Table IV. shows the number of subjects used. The university students were all men, the state normal school students were women and men, the latter composing about 6 percent.

Table IV

Showing Current Knowledge, Learning of False Authors, and True Authors of Unknown Literature (15 in Series)

Class of Literature	Class and No.[1] of Students	Nature of Response	Total Known	Mean	Median	Mode	Percentage
True Authors of Well Known Lit.	University[2] . 36	Remembering	390	10.8	11.5	13	72.20
	State Normal School.... 163	Rememberng	1,495	9.17	10.0	11	66.58
Less Known Lit.	University . . 38	Remembering	216	5.60	5.9	7	37.80
	S. N. S..... 139	Remembering	693	4.92	4.0	4	33.22
Unknown Literature..........	University . . 38	Immediate Recall	436	11.50	15	15	76.50
	S. N. S..... 178	Immediate Recall	2,323	13.00	14	15	87.00
False Authors of Well Known Lit.	University . . 36	Immediate Recall	391	10.86	12	15	72.40
	S. N. S..... 163	Immediate Recall	1,869	11.46	13	15	83.25
Less Known.....	University . . 38	Immediate Recall	515	13.55	15	15	90.35
	S. N. S..... 139	Immediate Recall	1,591	11.44	13	15	76.30
Well Known Lit.	University . . 35	Delayed Recall	258	7.37	7	8	49.14
	S. N. S..... 149	Delayed Recall	1,180	7.90	7	7	52.79
Less Known.....	University . . 34	Delayed Recall	226	6.35	6	5	44.31
	S. N. S..... 120	Delayed Recall	578	4.80	4	5	32.10
True Authors Unknown Lit...	University . . 38	Delayed Recall	387	10.1	11	11	67.8
	S. N. S..... 127	Delayed Recall	835	6.5	6	4	43.7

[1] Difference in number of students was due to absences from class.
[2] There were no college women in this part.

Results of Group Tests

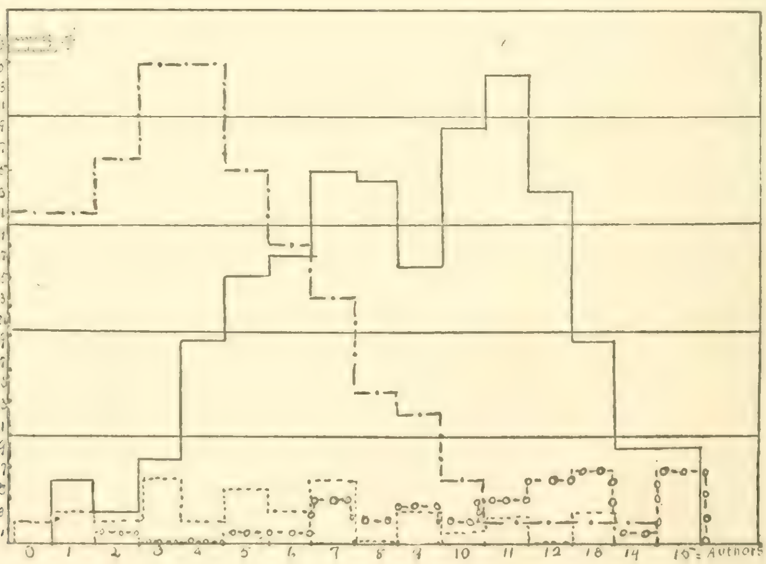
Inspection of Table I. shows the number of states considered 'familiar' by 45 college men and 100 college women. The alleged larger number of familiar states on the part of the women does not check with their current, or actual knowledge of state capitals. The college men show a 'current' knowledge of 30 percent and the college women of 27 percent of state capitals, both somewhat below our expectations.

But even with this comparatively small strength of bond between state and capital, it appears to have offered opposition to learning false capitals. The difference between the amount of learning of false capitals and of county seats shows this unmistakably. With the men, the difference is 8.5 percent and for the women 29.7 percent or an average of 19.1 percent in favor of the county seats.

But these are gross results derived from the responses of groups, and as such they fail to show whether the opposition to the false capital was offered by well-known, or less known, or barely known, or by capitals at the guessing stage of 'current' knowledge. Relying upon, as well as applying, the law of associative inhibition, one should claim that the well-known capitals offered the opposition. There is one way open, however, by which we may settle whether it is the well-known or the unknown capitals that exert the greater inhibition. From the data of current knowledge of capitals, two lists were derived, *known* and *unknown*, and it was quite possible to ascertain on a basis of ratios the relative opposition produced by the two sets of capitals. The ratio of the number of unrecalled false capitals to the number of states whose true capitals were known was computed, and, likewise, the ratio of the number of unrecalled false capitals to the number of states whose true capitals were unknown was reckoned. The ratios were computed alike for immediate and delayed recall and for recognition. For each form of response it was found that the ratio of the unlearned false to the known true capitals was less than the ratio of the unlearned false to the unknown true capitals. Table III. shows the numerical results from which the ratios were computed. They are expressed decimally.

So far as these facts go, they show that relatively strong associations between two mental elements offers less inhibition to the formation of a bond between either of the two and a third element, than a weaker bond between two such elements; and also that associations are more easily established between two equally new cognitive experiences than between old experiences void of prior connection with each other. The results of the check series, composed of county seats, referred to above, show this also.

GRAPH I, showing the frequency distribution of scores made by State Normal and University students in naming the authors of well-known and of less known literature.



Students	Number	Literature	Key to Graph
State Normal...........	336	Well-known	= ———————
State Normal...........	314	Less Known	= —-—-—-—-—
University.............	37	Well-known	= — o — o — o —
University.............	38	Less Known	= - - - - - - - - - - - - - -

As already indicated, the capitals classed here as 'unknown' must be, for the most part, *apparently*, and not actually, unknown. It is better to assume that the unknown is the *less known* capital, and that as such it offers greater inhibition to learning false capitals than the assumed well known.

The use of literature and authors enables one to pass from assumptions to positive knowledge about the learning material. When *well-known* literature was presented to 199 of the 374 subjects under test conditions, 63 percent of correct authors were named, while 177 of the same subjects under like conditions gave 34 percent of correct authors of less known literature. Diagram I. shows in graphic form the difference in degree of familiarity between the two types of literature. Table IV. shows the number and type of students and the mean, median, and mode, of each type in learning the three kinds of literature. The university students were juniors, the state normal students were juniors and seniors, about 80 percent the latter. These two distinct groups of students, working in quite different institutions, show a surprising uniformity of results. Their medians and modes very often agree and the differences are slight, their means differing only by one or two authors, and in one case only, by as much as 3.6, that occurring in unknown literature. Their respective percentages show differences ranging from 3.65 percent in the case of 'delayed recall' for false authors of well-known literature, to 14 percent occurring in 'immediate recall' of false authors in less known literature. This degree of uniformity of results seemed to justify combining them in Tables VI. and VII. It was found in the case of W.K. Lit. that the ratios of the unrecalled false to the respective known and unknown authors bore to each other the same relation as the similar ratios in the study with capitals, but in the L.K. Lit. the same ratios are reversed with but one exception, *viz.*, in immediate recall of false authors of L.K. Lit. by state normal students. Considering Table V., the results of the university students show a decided reversal in both immediate and in delayed recall of L.K. Lit. The ratios are, .1250 and .0762 in immediate recall, and .5797 and .3571 in delayed recall. That is, when a true author was known in L.K. Lit. greater inhibition was produced than if the true author had

Table V

INHIBITIONS OF FALSE AUTHORS IN LITERATURE

S. N. S. Students

Fifteen Well-known Authors

Nature of Response	Number Students	True Authors		False Authors Unrecalled	Ratios Known and Unknown True Authors to Unrecalled False Authors
Immediate recall	163	Known	1,495	304	.2073
	163	Unknown	950	225	.2368
Delayed recall	148	Known	1,375	590	.4290
	148	Unknown	860	424	.4930

J. H. U. Students

Immediate recall	36	Known	390	47	.1250
	36	Unknown	150	100	.6666
Delayed recall	35	Known	377	176	.4668
	35	Unknown	148	90	.6081

S. N. S. Students

Fifteen Less Known Authors

Nature of Response	Number Students	True Authors		False Authors Unrecalled	Ratios Known and Unknown True Authors to Unrecalled False Authors
Immediate recall	139	Known	693	150	.2164
	139	Unknown	1,392	323	.2327
Delayed recall	124	Known	593	356	.6003
	124	Unknown	1,267	588	.4561

J. H. U. Students

Immediate recall	38	Known	216	27	.1250
	38	Unknown	354	27	.0762
Delayed recall	34	Known	188	109	.5797
	34	Unknown	322	115	.3571

been unknown. The same thing was true of the normal school students in delayed recall. The less known authors

TABLE VI

INHIBITIONS TO FALSE AUTHORS IN WELL-KNOWN LITERATURE

Books	Total Known		I. Immediate Recall (199 Subjects) Unrecalled *False* Authors, when *True* Authors are: a. Known		b. Unknown		Total Known	II. Delayed Recall (183 Subjects) Unrecalled *False* Authors, when *True* Authors are: a. Known		b. Unknown	
	No.	Percent	Number	Percent	Number	Percent	Number	Number	Percent	Number	Percent
Faust	44	22.11	4	9.09	19	12.25	39	5	12.81	24	16.66
Gulliver's Travels	51	25.62	11	21.56	38	25.67	34	22	64.70	82	<55.03
Origin of Species	86	43.21	18	20.93	22	<19.46	80	34	42.50	54	52.42
Ancient Mariner	95	47.73	24	25.26	31	29.80	89	29	32.58	54	57.44
Les Miserables	96	48.24	12	12.50	19	18.43	88	35	39.77	36	<37.89
The Raven	126	63.31	24	19.04	9	<12.33	118	61	51.69	39	60.00
Snow Bound	128	64.32	27	21.09	22	30.98	118	62	52.54	39	60.00
Pilgrim's Progress	135	67.83	44	32.59	27	42.18	124	68	54.83	42	71.81
Treasure Island	146	73.36	43	29.45	18	33.96	135	66	48.88	28	58.33
Lady of the Lake	156	78.39	34	21.79	11	25.58	146	64	43.83	14	<37.83
David Copperfield	156	78.39	37	23.71	6	<13.95	148	58	39.18	11	<31.42
Paradise Lost	161	80.90	29	18.01	18	47.36	147	64	34.53	22	61.11
The Iliad	173	86.93	26	15.02	7	26.92	164	69	42.07	12	63.15
Hiawatha	182	91.46	45	24.72	13	76.47	168	99	58.92	11	73.33
Hamlet	193	96.98	30	15.54	5	83.33	179	74	41.34	4	100.00
Arith. Mean	128.43			20.68		33.24	118.47		44.61		55.76

To be read: Of 199 students 44 or 22.11 percent knew that Goethe wrote Faust.
Of 44 students who knew author of Faust, 4 or 9.09 percent did not recall the false author, Homer.
Of 155 students not knowing author of Faust, 19 or 12.25 percent did not recall the false author, Homer. (See page 277.)

To be read as same as Immediate Recall.

TABLE VII

INHIBITIONS TO FALSE AUTHORS IN LESS KNOWN LITERATURE

Books	Total Known		I. Immediate Recall (180 Subjects) Unrecalled *False* Authors, when *True* Authors are:				Total Number Known	II. Delayed Recall (159 Subjects) Unrecalled *False* Authors, when *True* Authors are:			
			a. Known		*b.* Unknown			*a.* Known		*b.* Unknown	
	No.	Percent	Number	Percent	Number	Percent		Number	Percent	Number	Percent
The Excursion	9	4.97	1	11.11	31	18.02	8	4	50.00	75	<49.66
The Ring and the Book	9	4.97	2	22.22	30	<17.44	13	3	23.76	73	50.00
The Man with the Hoe	19	10.49	1	5.26	28	17.27	10	5	50.00	60	<40.26
Rasselas	19	10.49	5	26.31	41	<23.29	17	9	52.94	104	73.23
Modern Painters	25	13.81	0	00.00	6	38.46	23	4	17.39	26	19.11
She Stoops to Conquer	36	19.88	4	11.11	32	22.06	42	8	19.04	59	50.42
The Skylark	42	23.20	5	11.90	22	15.82	36	20	55.55	58	<47.15
Tam O'Shanter	44	24.30	6	13.63	42	30.65	39	22	56.41	66	<55.00
Vision of Sir Launfal	71	39.23	13	18.30	29	26.36	61	36	59.01	48	<48.79
The Scarlet Letter	81	44.75	20	24.69	25	25.00	70	52	74.28	45	<51.86
Vanity Fair	91	50.27	24	25.38	14	<15.55	72	44	61.11	30	<34.48
The Inferno	106	58.56	16	15.09	19	25.33	93	50	53.76	38	57.57
The Jungle Book	113	62.98	27	23.89	17	25.00	102	65	63.72	33	<57.89
Sleepy Hollow	119	65.74	18	11.12	11	17.74	89	42	47.19	23	<32.85
Uncle Tom's Cabin	142	78.42	26	18.30	2	< 5.12	124	74	59.61	12	<48.00
Arith. Mean	61.73			17.88		21.54	53.27		49.58		47.55

To be read as Table VI.

are comparable in inhibiting force to the apparently unknown capitals. In other words, unknown capitals were really less known, and as such offered more inhibition than known. And, further, unknown authors in L.K. Lit. exert, apparently, a minimum of opposition, that is, they are approaching the condition of the unknown authors of the check material. According to this, the least unknown of the less known authors exert an inhibition similar to the best known of the well-known authors, *i.e.*, the extremes meet, so far as the inhibiting power is concerned, both are approaching the condition of the check material!

It appeared that a further analysis of the results might be made at this point by arranging the authors of the W.K. and L.K. Lit., in an array according to the number of authors known and then determining the amount of inhibiting of the individual authors.

The results of such an array and the computations made are expressed in Tables VI. and VII. The array of Table VI., while offering nothing new, confirms in a striking way the gross results already described. The arithmetical mean shows: (1) That inhibition is less in immediate than in delayed recall, (2) that the percentage of the unrecalled false authors is greater when the true author is unknown. But the arithmetical mean in delayed recall in Table VII., L.K. Lit., shows a reversal of that relation; *the inhibition of the known is greater than the inhibition of the unknown.* The arithmetical mean of the former is 49.58, and of the latter 47.55. The small difference between these two means is due to excessive variations in the 'unknown' column. The extent of the reversal is seen to a better advantage by comparing the percentages of the same author in the known and unknown columns. To aid in this the mathematical sign for 'less than' appears before those percentages in the unknown column that are less, respectively, than those in the known column. Ten of the fifteen authors have these signs in the unknown column of delayed recall in L.K. Lit. In a similar column of Table VI. four such signs appear.

To get another view of the relation between the degrees of knowledge and inhibition, the literary works of both W.K. and L.K. Lit. were arranged together in an array of thirty, according as their authors were known. The averages of the percentages of the first and fourth quartiles respectively and of the middle half of the percentages of the unrecalled false authors in both known and unknown columns were computed. These averages are as follows:

First quartile	Middle half	Fourth quartile
Known.......33.83%	Known.......53.77%	Known.......47.05%
Unknown.....43.06%	Unknown.....52.38%	Unknown.....59.20%

The first quartile of the known column, containing seven and one half literary works is composed of L.K. Lit. except Faust. The authors of this quartile were least known, and it had the fewest unrecalled false authors, 33.83 percent. The middle half, composed of 14 whole and two half works, or a total of 15, had an equal number of W.K. and L.K. Lit. There is an average of 53.77 percent of unrecalled false authors. The fourth quartile, having all W.K. Lit., except Uncle Tom's Cabin, has an average of 47.05 percent. The same quartile in the unknown column, containing the best known authors, has the highest percentage of unrecalled, 59.20 percent.

This quartile analysis of the percentages of the unrecalled false authors shows at least five sets of facts:

1. The least known of the L.K. Lit. offers no inhibition, but rather favors learning. The student had merely heard the names of the works, and that fact favored their association with even a false author (33.83 percent).

2. The entirely unknown of the L.K. Lit. (from the learners' standpoint) neither favors nor hinders learning; its status is neutral, and therefore comparable to that of new material. The percentage of the unrecalled is 43.06 percent, and this is very close to the percentage, 44.70 percent of the unrecalled true authors of the check material, Table IV.

3. The known authors of the L.K. Lit. and the unknown authors of the W.K. Lit. (middle half) offer decided opposition to making new associations (53.07 percent).

TABLE VIII

Showing Results of Reaction Time to True and False Associations in Geography and in Arithmetic

Subjects		True Capitals	False Association: Capitals		True Associations County Seats		True Associations Sums of Numbers	False Associations Sums of Numbers		Association Sums of Letters	
			Ser. I.	Ser. II.	Ser. I.	Ser. II.		Ser. V.	Ser. VI.	Ser. V.	Ser. VI.
	Mean	1.342	1.849		1.392		1.700	2.408	2.893	2.056	1.972
	Range	1.788	2.636		2.228		1.120	2.664	4.812	2.468	2.800
L.	P. E.	.0370	.0680		.0463		.1632	.1037	.1915	.0106	.0887
	Mean	1.580	8.860	5.312	5.730	1.412	1.585	3.370	3.392	2.293	4.560
	Range	2.472	9.536	18.728	10.448	1.576	2.968	11.252	2.340	6.384	8.800
M.	P. E.	.0511	.5736	1.1780	.0290	.0327	.0437	.3361	.1166	.1715	.2338
	Mean	1.712	2.060	1.928	1.520	1.904		3.000	5.164	5.288	2.236
	Range	2.940	26.324	10.484	1.832	1.940		3.312	13.180	25.748	2.376
R. H.	P. E.	.0490	.0867	.2537	.0470	.0512		.1087	.0434	.0793	.0900
	Mean	1.692	3.300	3.924	4.512	1.568	1.464	4.776	3.804	1.912	3.456
	Range	3.164	8.620	5.126	4.236	0.248	.740	12.656	3.008	2.696	4.784
B.	P. E.	.0666	.0227	.0926	.1982	.0585	.0199	.0458	.1900	.0830	.1619
	Mean	1.280	2.550	2.088	1.848	3.704	1.300	2.372	2.544	3.140	2.696
	Range	2.080	4.388	2.932	1.224	11.764	1.503	3.516	2.460	5.780	2.944
W.	P. E.	.0565	.1933	.1339	.0571	.2433	.0269	.1038	.0999	.2350	.1304
	Mean	2.316	5.012	2.228	6.908	1.328	1.628	6.448	5.864	2.300	2.744
	Range	5.276	7.652	4.132	9.120	2.116	1.516	7.268	16.744	1.876	2.096
C.	P. E.	.0301	.2591	.0919	.0770	.0489	.0255	.0588	.0519	.1133	.0722
	Mean	1.232	5.032		1.516			2.764	3.040	1.604	1.980
	Range	2.924	27.324		2.752			5.840	2.964	1.864	2.256
E. H.	P. E.	.0348	.0665		.0710			.1266	.0955	.0573	.0912
Mean of total:											
	Means	1.593	4.095	3.096	3.346	1.983	1.535	3.591	3.814	2.656	2.806
	Ranges	2.949	12.350	8.298	4.548	3.528	1.569	8.072	6.501	6.688	3.722
	P. E.'s	.0464	.1813	.3500	.0750	.0865	.0558	.1287	.1126	.1071	.1240

4. The unknown of the best known (fourth quartile) of W.K. Lit. offers a maximum of inhibition (59.20 percent).

5. The known of the best known of W.K. Lit. offers comparatively little inhibition, if we measure it by the difference between the percentage of unrecalled false authors, 47.05 percent, and the unrecalled authors of the check series 44.70 percent, a difference of only 2.33 percent.

B. Procedure with Individual Subjects

Twelve subjects were used, seven with the false geography, and seven with the false authors, two serving in both. Four of the geography group were college graduates, two of whom were psychologists. The remaining four were undergraduates and had not studied psychology. *L* and *C* (Table VIII.) knew the purpose of the experiment.

(*a*) The same geography material used in the group experiment was used here. The subject's current knowledge of the captials and familiarity with the states used in the experiment were first ascertained. Then his normal constrained association time for giving true capitals of states was determined by the Johns Hopkins chronoscope. Twenty states were used for this purpose and also twenty foreign countries. While the current knowledge of foreign capitals was much higher than that of the states, the association reactions showed 400^{σ} shorter for the states.

To subjects unfamiliar with making verbal responses to a chronoscope, preliminary practice was given. A period of 45 minutes or more was required for learning both the false and check material. The learning time was 15 minutes each for the capitals and the county seats, with an interval of 10 minutes between. Immediate recall was made in both cases by writing as the name of state or county was pronounced. The pronouncing order was different from the learning order. After an interval of 48 hours both series were reviewed, five minutes being devoted to each. At the end of the ten minutes review the subjects' reaction time was determined by the chronoscope. The names of the states and of the counties were pronounced by the experimenter in an irregular order, the subject responding verbally to each with the name of its appropriate false capital or county seat as the case might be. The Johns Hopkins chronoscope made it possible to get constrained association time readily and free from distractions. Two sets of countries were used, Series I., European, and Series II., States of the Union. Sample reaction times are shown here in Series I. and Series II.

SERIES I. *Groups (a) and (b) Combined for Comparison*

Country and County Alternating	Capital or County Seat	Reaction Time
Italy	Lisbon	554
Chester	Henderson	696
Germany	Stockholm	245
Lamar	Vernon	270
Spain	Warsaw	605
Early	Blakeley	639
Turkey	Madrid	608
Meigs	Pomeroy	392
Austria	Brussels	236
Tomba	Lyons	358
France	Tokio	496
Bradford	Stock	518
Poland	Rio de Janeiro	656
Hernando	Ingham	443
Brazil	Paris	395
Turner	Ashburn	378
Japan	Edinburgh	276
Franklin	Yeager	625
Ireland	Berne	530
Hardeman	Bolivar	530
England	Mexico City	508
Bond	Greeneville	750
Switzerland	London	*
Gadsden	Quincy	267
Belgium	Vienna	2620
Terrill	Dawson	323
Portugal	Berlin	375
Obion	Union City	814
Mexico	Dublin	288
Kanabec	Morrow	*

	Capitals	County Seat
Average	= 513.7σ	500σ
Median	= 502.	481

*No record taken.

SERIES II. *Arranged for Testing*

State and County Alternating	Capital or County Seat	Reaction Time
Pike	Magnolia	380
Virginia	Montgomery	1429
Copiah	Hazelhurst	238
Idaho	Little Rock	275
Pearl	Poplarville	114
Kentucky	Madison	488
Webster	Walthall	297
Nebraska	Richmond	372
Wilcox	Camden	575
Tennessee	Sacramento	148
Alcorn	Corinth	430
Rhode Island	Denver	466
Barbour	Clayton	362
Georgia	Tallahassee	1033
Quitman	Belan	263
Wisconsin	Indianapolis	370
Sharkey	Rolling-Fork	497
Indiana	Prescott	785
Tate	Senatobia	437
California	Springfield	333
Benton	Ashland	619
Arizona	Dover	154
Tippah	Ripley	176
South Carolina	Hartford	1124
Leflore	Greenwood	271
Missouri	Montpelier	370
Homes	Lexington	238
New Hampshire	Olympia	312
Hinds	Raymond	90
New Jersey	Charleston	125

	Capitals	County Seat
Average	= 568.9σ	332σ
Median	= 370.	271σ

The combined results are given in Table VIII. Series II., composed of states, was regarded by the subjects as more difficult, but the general average of the association times is shorter than that of Series I., requiring false capitals for European countries. Usually the difference might be attributed to practice, and for some of the subjects that appears to be the case. *W* and *R. H.* learned Series I. first, and had a shorter association time than Series II. *M*, *B*, and *C*, learned Series II. first, and here practice effects appeared only with *B*. That is, practice effects are evident in three, and not apparent in two cases. With two subjects, *B* and *C*, the time of reaction in one of the check series is longer. They attribute it to the confusion arising from the names of certain county seats ending in 'ville.' In all other instances a shorter time was registered.

The manner of learning in both the test and check series followed the same lines. (1) Mnemonic devices where employed followed the same lines, *e.g.*, Kentucky-Mad-i-son. "Kentucky makes whiskey, whiskey makes madison." (2) Ingenious intermediaries were resorted to. (3) 'Straight-memory' as termed by the subject.

(*b*) *The Association of False Sums of Numbers.*—In general the procedure with this method followed that observed with the false capitals. Each series had only ten pairs, as given here.

Series V	Series VI	Check Series V
2 + 2 = 13	10 + 13 = 27	*L* + *P* = 8
1 + 3 = 2	9 + 14 = 20	*R* + *O* = 7
3 + 7 = 12	7 + 15 = 43	*G* + *M* = 12
4 + 5 = 7	11 + 12 = 19	*N* + *K* = 11
6 + 2 = 9	13 + 16 = 30	*Q* + *L* = 17
5 + 3 = 6	8 + 17 = 22	*B* + *E* = 13
8 + 3 = 23	12 + 14 = 28	*O* + *V* = 14
4 + 2 = 5	14 + 8 = 18	*K* + *R* = 19
2 + 7 = 15	6 + 9 = 16	*T* + *J* = 16
8 + 4 = 10	15 + 6 = 17	*D* + *C* = 21

Ten minutes were devoted to learning. Immediate recall was done by both writing and pronouncing the false or arbitrary associate as the experimenter pronounced the numbers or the letters of the series in an irregular order. After an interval of 48 hours a review of five minutes, with some subjects, eight minutes, was given. In the latter case, the same time was extended to the letter series, even though it was not needed. Twenty pairs of numbers and their true sums were used to ascertain the normal association time. This time based here upon 100 responses, is 1.535^{σ} for true sums, and true capitals, based upon 140 responses, is 1.593^{σ}.

Results

Of the two series of numbers, V. and VI., the latter was considered the more difficult. This was due to the fact that Series V. contained all the simpler and smaller number combinations, and to avoid similarities in the two series, larger numbers had to be used. Although V. was given first, the effect of practice on VI. was obscured by its difficulty. Associating letters and their sums was considered less difficult than numbers and false sums by five of the subjects. The reason assigned was the difficulty in finding familiar and extra associations to serve as connecting bonds between numbers and sums. Two subjects, *M* and *W*, claimed that letters and sums were more difficult. This checks with results of both *W* and *M* in Series VI. *W*, who had a good place memory, memorized the pairs as they occurred consecutively in the series. He adhered to it throughout, despite the fact that the order was always changed in recall. The subjects, as a rule, used mnemonic devices and useful intermediaries more freely than in learning false capitals. For example, $12 + 14 = 28$: the second number is one half of the false sum 28. $U + H = 31$, because *U* (ta) *H* is the 31st state! etc. Attention has been called to the fact that the normal constrained association time for true capitals and that for true sums were practically the same. The inhibiting power of the false capital is somewhat less, however. A false capital increases the normal association time by 2.002″, and a false sum increases the normal time by 2.168″. (Derived from Table VIII.)

(*c*) *The Association of False Authors in Literature.*—The same authors, arranged in the same series, were used here as with the groups; and, in addition, an extra series of unknown authors was prepared for use with the L.K. Lit.

Check Series I		Check Series II	
Kind Hart's Dream	Henry Chettle	Revoked Vengeance	Heard
Anne Hereford	Wood	The Country Ball	Praed
The Doome	Batman, etc.	Shadows on the Wall	Hewitt, etc.

The main purpose in using this material with individuals was to see if the results would check with those of the groups, showing that the less known literature was more inhibitive than the well known. The comparative strength of the two types, *i.e.*, W.K. Lit. and L.K. Lit., was sought in two ways: (1) By comparing the difference between the association time between W.K. Lit. and its check series (unk. lit.), with the difference that occurred between L.K. Lit. and its check series. (2) By combining the W.K. and L.K. types into a single series of 30, using the same false authors and determining again their association time. This latter procedure insured relatively uniform conditions in responding to the two series, and gave results that permitted direct comparison.

The limited current knowledge of both types of literature made it impractical to secure normal association time. It was attempted with uncertain results for the W.K. series, and to have required learning of the true authors would have changed the procedure from that employed with the groups. The seven subjects knew, on an average, 66 percent of the authors of the W.K. and 33 percent of the authors of the L.K. series. The entire procedure covered six periods given here in schematic form.

Period I	Period II	Period III
Learning W.K. paired with check series	Testing after 48 hours	Learning (1 wk. later than Period II) L.K. paired with check series.
Period IV	**Period V**	**Period VI**
Testing after 48 hours	(Wk. after Per. IV) W.K. paired with L.K. series	Testing after 48 hours.

Table IX

Showing Reaction Time to False Authors of Well and Less Known Literature and True Authors of Unknown Literature

Kind of Measure	When Well-known Literature Series Is Presented First (4 Subjects)					
	W.K. Paired with Unk. Lit.		L.K. Paired with Unk. Lit.		W.K. Paired with L.K. Lit.	
	Well-known	Unknown	Less Known	Unknown	Well-known	Less Known
Mean	3.084	2.492	2.520	1.728	2.628	1.992
Mean	1.576	1.684	2.088	1.624	1.788	1.944
Mean	2.252	1.996	1.324	1.164	4.768	2.008
Mean	7.128	2.816	4.244	4.792	1.228	4.956
Av. of Mean	3.510	2.247	2.544	2.327	2.603	2.725
Range	4.368	6.940	8.648	2.704	6.700	5.032
Range	2.556	2.144	5.204	1.692	1.780	7.120
Range	3.440	5.920	1.940	1.440	3.052	7.260
Range	21.168	6.664	6.520	4.720	9.920	8.636
Av. of Range	7.883	5.417	5.578	2.639	5.363	7.012
Median	3.120	2.080	1.800	1.364	1.790	1.734
Median	1.430	1.560	1.700	1.640	1.000	1.260
Median	2.120	1.440	1.248	1.064	1.640	1.460
Median	3.780	2.134	3.544	3.540	3.740	4.168
Av. of Median	2.612	1.803	2.073	1.902	2.042	2.155
P. E.	0.1126	0.1141	0.2531	0.0727	0.1701	0.1102
P. E.	0.0522	0.0463	0.1163	0.0444	0.0444	0.0187
P. E.	0.0964	0.1223	0.0390	0.0381	0.0722	0.1494
P. E.	0.1675	0.0479	0.0575	0.0306	0.0277	0.2245
Av. of P. E.'s	0.1071	0.0901	0.1164	0.0464	0.0786	0.1257
	When Less Known Literature Series Is Presented First (3 Subjects)					
Mean	2.088	2.072	2.152	1.332	1.740	1.524
Mean	2.224	1.448	1.812	1.464	1.732	1.460
Mean	2.332	1.568	2.681	2.092	1.504	1.224
Av. of Mean	2.214	1.696	2.216	1.629	1.658	1.402
Range	3.688	2.980	5.220	1.240	2.700	2.992
Range	6.268	1.540	3.312	1.532	5.588	2.404
Range	4.880	1.304	4.104	3.860	3.060	1.480
Av. of Range	4.942	1.941	4.212	2.210	3.782	2.292
Median	1.692	1.480	1.540	1.372	1.580	1.400
Median	1.680	1.660	1.816	1.540	1.420	1.180
Median	1.621	1.616	1.440	1.340	1.380	1.216
Av. of Median	1.665	1.585	1.598	1.417	1.460	1.265

TABLE IX—*Continued.*

Kind of Measure	When Less Known Literature Series Is Presented First (3 Subjects)					
	W.K. Paired with Unk. Lit.		L.K. Paired with Unk. Lit.		W.K. Paired with L.K. Lit.	
	Well-known	Unknown	Less Known	Unknown	Well-known	Less Known
P. E..............	0.1007	0.0846	0.1407	0.0249	0.7030	0.0653
P. E..............	0.0444	0.0364	0.0724	0.0342	0.1164	0.0578
P. E..............	0.1339	0.0327	0.0883	0.0900	0.0582	0.0324
Av. of P. E.'s......	0.0930	0.0512	0.1004	0.0497	0.0816	0.0518

Four subjects learned W.K. Lit. first, and three, L.K. Lit. first. The results are given in Table IX. Differences between the averages of the means for the several paired types follow, also P.E.'s:

I. When W.K. Lit. was Presented First:

	(*a*) W.K. paired with Unk.	(*b*) L.K. paired with Unk.	(*c*) W.K. paired with L.K.
Difference of Means..	1.263″ excess W.K.	0.217″ excess L.K.	0.122″ excess L.K.
Difference of P.E.'s...	0.017″ excess W.K.	0.070″ excess L.K.	0.0471″ excess L.K.

(To be read: The average association time for responding with false authors of W.K. Lit. exceeds the average association time for responding with true authors to unknown literature by 1.263″.)

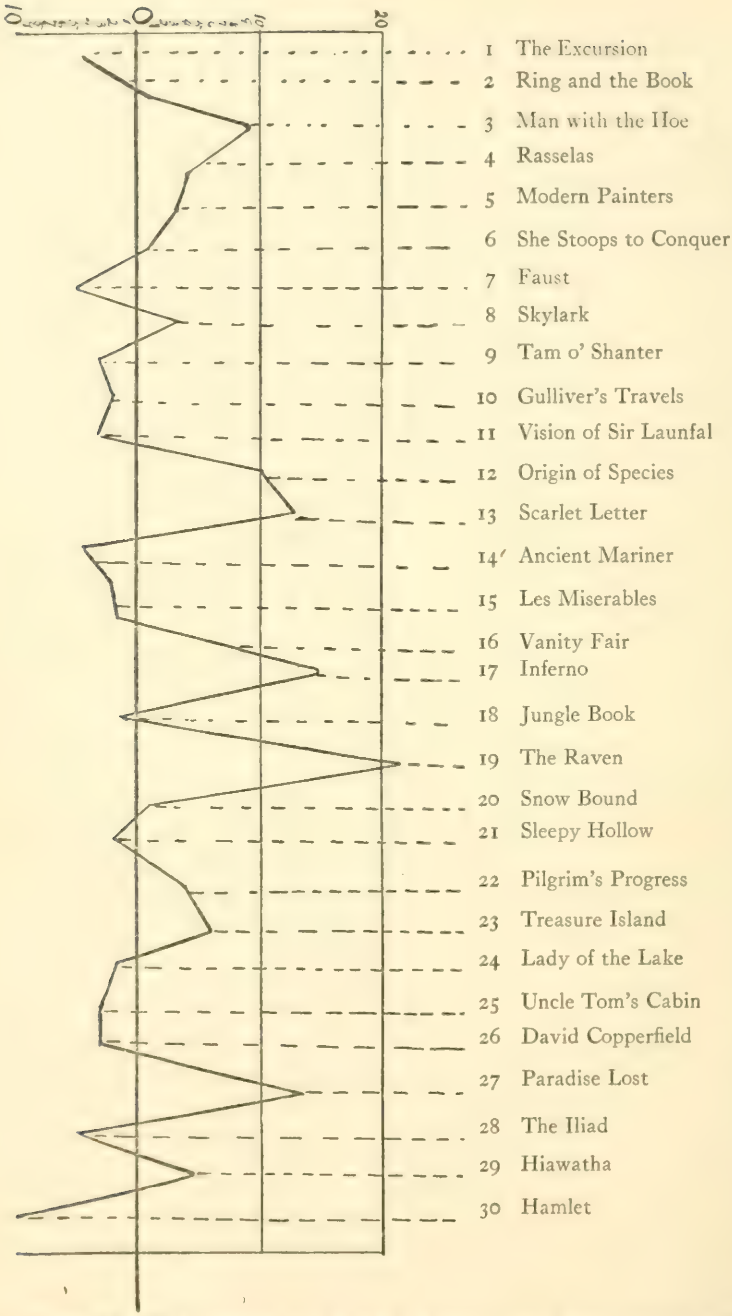
II. When L.K. Lit. was Presented First:

	(*a*) W.K. paired with Unk.	(*b*) L.K. paired with Unk.	(*c*) W.K. paired with L.K.
Difference of Means.	0.518″ excess W.K.	0.527″ excess L.K.	0.256″ excess W.K.
Difference of P.E.'s ..	0.0418″ excess W.K.	0.050″ excess L.K.	0.0298″ excess W.K.

This partial synopsis of Table IX. shows that when W.K. series is presented *first*, the difference in association time between W.K. and its check series is greater than that between L.K. and its check series; but when L.K. and W.K. are combined in the same series, L.K. shows the excess (see (*c*) above). When the L.K. series is presented first, results are reversed in the third combination (*c*), and W.K. means or averages are greater than those of L.K. The excess of L.K. over Unk. is more than doubled, while that of W.K. is less than half, this series giving the shorter association time as expressed by both means and medians. Whatever amount of inhibition may be produced under the conditions, it is obscured by the effect of practice; and the question as to the relative inhibiting effects of W.K. and L.K. Lit. is left in doubt so far as this method of comparison is concerned.

A simpler and more direct method of comparison is found by using the quartile distribution of the association times. Recourse was had to the array of W.K. and L.K. established by the results obtained by working with groups, in which the two were combined into a single series according to the number of authors known. The average association time of seven subjects, responding with a false author to his associated literary work, was computed for the entire series of 30 literary works in the array. *The average in each case is based upon the responses of seven subjects.* The Excursion was first in the array and gave an average association time for seven subjects of .992″. Tam o' Shanter came ninth and gave 2.920″, and Hamlet came thirtieth or last with 0.892″. Then the averages of the association times for the first and fourth quartiles, and the middle half of the array were computed; the first quartile gave an average of 1.572″, the middle half 1.760″. While the differences in association time may be considered relatively small, I think they are fully large enough to be significant, and to afford unmistakable support to the results obtained from subjects taken in groups. Least known offer a *minimum* of inhibition, less, fairly or half known authors produce a *maximum*, while well-known authors give a *medium* inhibition. Or in general, the second type shows, that a maximum amount of inhibition is exerted in forming new bonds of connections between a new and an old experience when the latter has weak bonds with other experiences, while the third indicates that a medium amount is offered in forming new bonds of connections between an old and a new experience when the former has firm connections with other experiences.

These several degrees of inhibition are best expressed in Graph II. Here the amount exerted is measured by the difference in the association time for responding with a false author as shown by the mean for the seven individual subjects



GRAPH II. The distance of the graph above o line shows relative inhibiting strength of true author. o line is the normal association time for responding to true author of unknown literature.

and the mean of the association time for responding with a false author as shown by the mean for the seven individual subjects and the mean of the association for unknown authors based upon the grand total of responses of the seven subjects to the 30 unknown literary works. This latter association time was 1.975. It serves as the base of reference in Graph II. and is marked 'o.' If the association time for a false author was less than 1.975, the difference is indicated by the distance below the zero line, *e.g.*, the mean time for giving the false author of 'Excursion' is 1.528, or less than that of the mean for the unknown by 0.447. It is indicated in the graph as over four units below the base line. If the association time of the false author exceeds that of the mean of the unknown, the difference was regarded by that much as inhibition and is expressed by distance above the base line. The false author of Skylark gave an association time of 2.316 or greater than the mean of the unknown by .341.

Both the W.K. Lit. and L.K. Lit. are combined into an array as previously described and placed above the graph so that the name of each literary work falls directly over the point that measures the extent of its variation above or below the base line.

The graph affords, perhaps, the best device for showing the extent to which the results obtained from the individuals check those of the groups. The factors upon which the graph is based are too complex, however, to justify an explanation of its several irregularities, were it possible to do so.

Summary and Discussion

1. The results make it clear that the comparative difficulty of both State Capitals and Authors should have been determined prior to the investigation by testing at least 100 representative college students. This applies particularly to the authors. Had this been done, Faust, Gulliver's Travels, Origin of Species of the W.K. Lit. would doubtless have changed places with The Jungle Book, Sleepy Hollow, and Uncle Tom's Cabin in L.K. Lit. While this arrangement would hardly have changed the direction of the results, it would have increased their decisiveness.

2. The investigation shows that associative inhibition is an active principle in learning meaningful material, acting adversely for the most part, and that the amount of its action is determined by the relative strength of the association between the two elements involved.

3. If we measure the inhibiting strength offered by an author and his book, or any other associative material, according to the percentage of *group recall* or 'current knowledge,' then the degrees of strength can be expressed numerically. And from the present results we may obtain the following degrees of strength:

(*a*) If the bond is so weak as to be practically inoperative, having a recall power from 1 percent to 10 percent, and if one of the elements be somewhat familiar, then connecting the latter with a third element is favored by the associative condition (Excursion in Table VII. and Graph II.).

(*b*) If the connecting bonds give a recalling power of 15 percent to 40 percent, the inhibiting power is relatively small (She Stoops to Conquer, Table VII., also Graph II.).

(*c*) If the connecting bonds give a recalling power of 45 percent to 70 percent, the inhibiting strength approaches a maximum degree. Still, "a little learning is a dangerous thing."

(*d*) If the connecting bonds give a recalling power of 75 percent to 100 percent, the inhibition is of medium strength and in some cases may pass from an inhibiting to a facilitating rôle in learning (see Iliad, Hiawatha, and Hamlet, in Table VI. and Graph II.).

The two degrees of strength described in (*c*) and (*d*) are illustrated by new and old postal clerks in re-learning the distribution of a mail route after it has been broken up and re-formed into a new distribution. The clerk old in service re-learns and uses the new arrangement with greater ease and certainty than the new clerk of small experience. Of course the groups upon which the percentage of recall power is based should be homogeneous. The groups of this investigation have already been described.

4. The principle of inhibition in the higher association

processes grounds on the relative strength of the reacting complexes involved, or, in neural terms, it appears to depend on the difference of preparedness for discharge that may exist between rival neural centers. The functioning of the advantaged center, for whatever reason, precludes the functioning of its rival. A thoroughgoing theory of inhibition must account for (1) The difference in preparedness between the rival centers, so-called, (2) It must account for neural centers becoming rivals, or opposed to each other. Is this more or less paired opposition acquired or innate, or both? The present study indicates that it may be acquired, when involving high association processes.

The bibliography appended here contains only the works consulted in this investigation.

1. Bair, Joseph H. The Practice Curve: A Study in the Formation of Habits. Psychol. Rev. Monog. Suppl. No. 19, 1902, pp. 70.
2. Bergstrom, John A. Experiments Upon Physiological Memory, *Amer. Jour. of Psychol.*, Vol. 5, 1893, pp. 356–359.
3. Bergstrom, John A. The Relation of the Interference to the Practice Effect of an Association. *Amer. Jour. of Psychol.*, Vol. 6, 1895, pp. 433–442.
4. Culler, A. J. Interference and Adaptability. *Arch. of Psychology*, No. 24, 1912, pp. 80.
5. Curtis, H. S. Inhibition. *Ped. Sem.*, Vol. 16, pp. 65–113.
6. DeCamp, J. Edgar. A Study of Retroactive Inhibition, Psychol. Monog. Vol. 19, No. 4, Aug. 1915.
7. Kline, L. W. and Owens, W. A. The Learning Process. *Psychol. Rev.*, Vol. 20, 1913, pp. 237–244.
8. Müller, G. E. u. Schumann. Experimentelle Beiträge zu Untersuching des Gedächtnisses. In der *Zeitschrift f. Psychologie*, 6 Bd., 1894.
9. Müller, G. E. and Pilzecker, A. Experimentelle Beiträge zu Lehre von Gedachtniss. *Zeitschrift für Psychologie*, erganzungs, Bd. I, 1900.
10. Münsterberg, Hugo. Psychotherapy, Chap. 5, pp. 85–124, 1909.
11. Shepard, J. F. and Fogelsangu, H. M. Studies in Association and Inhibition. *Psychol. Rev.*, Vol. XX., No. 4.
12. Stratton, G. M. Retroactive Hypermnesia. *Psychol. Rev.*, Vol. 26, pp. 474–475, 1919.
13. Swift, Edgar J. Psychology and the Day's Work, 1918.
14. Tolman, E. D. Retroactive Inhibition as Affected by Conditions of Learning. Psychol. Rev. Monog. Suppl., Vol. XXV., No. 1, 1918.

EFFECTS OF PRACTICE UPON THE SCORES AND PREDICTIVE VALUE OF THE ALPHA INTELLIGENCE EXAMINATION

BY FLORENCE RICHARDSON AND EDWARD S. ROBINSON

The University of Chicago

I. Subjects, Materials and Methods

The material upon which this report is based was derived from the examinations given a group of students who were used as controls in a study, shortly to be published, on the effects of loss of sleep.

Three forms of the alpha examination in the order 5, 7, and 9,[1] were given during a forenoon class period on three consecutive days, March 2 to 4 inclusive. These examinations were chosen on the assumption that they are well equated in difficulty, and that the scores of the different forms are comparable.[2]

The examinations were given on each day and scored in strict accordance with the instructions in use in the Army as

[1] Dunlap and Snyder (Jour. of Exp. Psych., Oct., 1920, Vol. III., pages 396–403) used forms of the alpha examination in the order 6, 8, 5, and 9, at intervals of 'approximately three weeks.'

[2] Since this paper was written the fuller account of the army results from the alpha examination has come to hand (*Memoirs of the National Academy of Sciences*, 1921, Vol. XV.). Data are presented here (p. 659, Table 174, and p. 660, Tables 176 and 177) which show the relative difficulty of these forms for random samplings of the white draft. The differences in the difficulty of the different forms according to these data are hardly great enough to alter to an important extent the results presented in this paper.

given by Yoakum and Yerkes.[1] The group was composed of thirty-nine college students, thirty-seven men and two women, most of them of sophomore standing and of an average age of 20.2 years. Of these, one had taken the alpha examination several years ago; he ranked thirty-first among our subjects on the first day's trial with a score of 144. Twenty-seven had taken the Thorndike Intelligence Tests for High School Graduates and College Students about sixteen months earlier; their average score on the first day's test was 160.9 as compared with the general average of 160.0. Six had taken the Thurstone Intelligence Test for High School Graduates, and their average score for the first day's test was 161.8. Six had taken no such test previously and their average score on the first day was 162.6. It seems likely that possible differences due to previous contacts with such test materials had little positive significance as practice on the scores of the first day.

The effects of practice are considered from the standpoint of *scores*, which represents the amount of credit given according to the army method of scoring the *right* responses, and of *attempts*, which represents the total number of items responded to, either correctly or incorrectly. Not enough work has been published comparing these two methods to justify the elimination of either in estimating the relation of one examination to another.

The statistical analysis of the results gives information of value to those who make practical use of the tests in college administration. It is interesting also to consider these scores as an example of practice in a complex function, and of the mechanisms by means of which positive and negative transfer operate under these specific conditions.

II. Absolute and Relative Improvement

Table II. presents the average scores for the three days of the experiment and also the relative achievements on days 2 and 3 as compared with day 1. According to the combined scores of all eight tests, day 2 shows a marked improvement over day 1, and day 3 a less marked improvement

[1] 'Army Mental Tests,' Yoakum and Yerkes, 1920, pp. 53–79.

Table I

Examination Records of Individual Subjects

S—Score
A—Attempts

Subject	Day 1		Day 2		Day 3	
	Form 5		Form 7		Form 9	
	S	*A*	*S*	*A*	*S*	*A*
1	198	207	200	212	203	212
2	193	205	201	211	194	212
3	191	203	181	206	196	201
4	187	208	182	201	185	208
5	180	194	171	196	182	207
6	178	199	182	194	179	197
7	176	191	193	209	200	212
8	176	183	177	194	192	202
9	176	201	185	200	182	203
10	175	196	173	202	168	198
11	174	200	177	203	163	201
12	173	188	176	194	180	206
13	169	192	180	200	174	207
14	168	197	186	208	187	208
15	166	184	187	199	182	200
16	165	194	178	200	183	198
17	165	190	159	194	160	201
18	165	193	170	189	177	201
19	161	173	152	167	162	192
20	159	195	168	199	166	204
21	159	185	157	198	176	198
22	159	199	174	208	167	199
23	158	185	178	198	176	201
24	158	171	171	190	178	202
25	157	166	166	189	161	181
26	156	204	158	205	161	204
27	155	179	162	192	163	192
28	153	190	180	199	175	203
29	151	200	165	207	177	210
30	151	191	158	201	156	201
31	148	163	169	197	164	193
32	147	167	162	185	163	194
33	147	159	173	182	175	185
34	144	168	152	191	151	189
35	140	165	157	182	153	180
36	138	159	159	190	153	189
37	131	170	161	206	149	201
38	130	166	160	187	155	188
39	102	131	111	141	122	155

Table II

Average Test Scores and Relative Achievement on Successive Days

Combined Records (All tests)

	Day 1	Day 2	Day 3
Average and S. D.	161.00 ± 18.63	170.51 ± 15.9	171.54 ± 15.2
Relative Achievement	100%	105.9%	106.5%

TABLE II—*Continued*

Test 1, Oral Directions

	Day 1	Day 2	Day 3
Average and S. D.	10.13 ± 1.35	11.46 ± 0.62	11.49 ± 0.70
Relative Achievement	100%	113.1%	113.4%

Test 2, Arithmetical Problems

	Day 1	Day 2	Day 3
Average and S. D.	12.79 ± 2.50	14.62 ± 2.35	15.67 ± 2.73
Relative Achievement	100%	114.3%	122.5%

Test 3, Practical Judgment

	Day 1	Day 2	Day 3
Average and S. D.	12.36 ± 2.67	13.49 ± 2.08	14.21 ± 1.21
Relative Achievement	100%	109.1%	114.9%

Test 4, Synonym-Antonym

	Day 1	Day 2	Day 3
Average and S. D.	30.41 ± 5.91	29.74 ± 5.54	30.79 ± 5.02
Relative Achievement	100%	97.8%	101.2%

Test 5, Disarranged Sentences

	Day 1	Day 2	Day 3
Average and S. D.	19.08 ± 3.48	21.51 ± 3.26	19.69 ± 3.29
Relative Achievement	100%	112.7%	103.2%

Test 6, Number Series Completion

	Day 1	Day 2	Day 3
Average and S. D.	12.79 ± 2.49	13.85 ± 2.80	15.36 ± 2.69
Relative Achievement	100%	108.3%	120.1%

Test 7, Analogies

	Day 1	Day 2	Day 3
Average and S. D.	29.82 ± 5.46	34.95 ± 3.18	32.90 ± 4.83
Relative Achievement	100%	117.2%	110.3%

Test 8, Information

	Day 1	Day 2	Day 3
Average and S. D.	33.69 ± 3.7	30.92 ± 4.1	31.44 ± 3.9
Relative Achievement	100%	91.8%	93.3%

Table III

AVERAGE ATTEMPTS AND RELATIVE ACHIEVEMENT ON SUCCESSIVE DAYS

Combined Records (All Tests)

	Day 1	Day 2	Day 3
Average and S. D.	184.89 ± 16.82	195.54 ± 11.6	198.33 ± 10.44
Relative Achievement	100%	105.8%	107.3%

Test 1, Oral Directions

	Day 1	Day 2	Day 3
Average and S. D.	12.0 ± 0	12.0 ± 0	12.0 ± 0
Relative Achievement	100%	100%	100%

Test 2, Arithmetical Problems

	Day 1	Day 2	Day 3
Average and S. D.	15.13 ± 2.26	16.36 ± 1.93	17.18 ± 2.28
Relative Achievement	100%	108.1%	113.6%

Test 3, Practical Judgment

	Day 1	Day 2	Day 3
Average and S. D.	13.31 ± 2.50	13.74 ± 2.15	15.44 ± 1.18
Relative Achievement	100%	103.2%	116.0%

Test 4, Synonym-Antonym

	Day 1	Day 2	Day 3
Average and S. D.	35.54 ± 5.58	36.18 ± 4.06	37.74 ± 3.52
Relative Achievement	100%	101.8%	106.2%

Test 5, Disarranged Sentences

	Day 1	Day 2	Day 3
Average and S. D.	21.92 ± 2.50	23.62 ± 1.05	23.31 ± 1.40
Relative Achievement	100%	107.7%	106.3%

Test 6, Number Series Completion

	Day 1	Day 2	Day 3
Average and S. D.	14.87 ± 2.58	16.08 ± 2.62	17.05 ± 2.18
Relative Achievement	100%	108.1%	114.7%

TABLE III—*Continued*

Test 7, Analogies

	Day 1	Day 2	Day 3
Average and S. D.	34.00 ± 4.90	38.69 ± 2.80	36.31 ± 3.76
Relative Achievement	100%	113.8%	106.8%

Test 8, Information

	Day 1	Day 2	Day 3
Average and S. D.	38.13 ± 4.08	38.90 ± 3.96	39.31 ± 1.60
Relative Achievement	100%	102.0%	103.1%

over day 2, *i.e.*, progress due to practice is, within the limits studied, negatively accelerated.[1,2] Variability as measured by the standard deviation decreases roughly in inverse proportion to improvement.

[1] Yoakum and Yerkes (*op. cit.*, p. 50) state that the average gain in a repeated alpha examination is approximately 8 points (raw score). If this gain refers to the gain of a second testing over a first, it is approximately substantiated by our results which show a gain of 9.47 points with this sort and amount of practice.

[2] If we should correct the average scores for these three days in light of the data presented in the *Memoirs*, Vol. XV., our facts would be little altered. No accurate correction can be derived for our results, however, from the army data which are based upon the random samplings of the white draft. Whereas the average scores of that group in the different forms range from 60.85 to 66.81, the average score for our group before practice begins to take effect is 161.07. It would hardly be legitimate to conclude that differences in the difficulty of different forms as shown by the performances of a group whose average performance is 60–67 would hold for a group whose average performance is 160 or better.

In order to estimate the relative difficulty of forms 5, 7, and 9 (the forms employed in our study) in the upper ranges of intelligence, we computed the average performances with these forms by those of the white draft whose scores were 100 or better (our lowest score was 102). According to the army data from the complete samplings, form 7 is 2.44 points more difficult than form 5 and form 9 is 3.47 points more difficult than form 7, while according to our selections of the more intelligent members of the white draft, form 7 is 1.00 easier than form 5, and form 9 is 4.20 more difficult than form 7. While our sampling of the more intelligent members of the white draft is done in quite arbitrary fashion, it is perhaps fair to judge from it that forms 5 and 7 are practically of equal difficulty for groups of a high degree of intelligence and that form 9 is about 4 to 5 points more difficult than form 7.

It is probable, then, that the amount of improvement from day 1, form 5, to day 2, form 7, as shown in Table II. is much more a matter of practice than of the relative difficulties of the two forms. Since form 9, which was employed on day 3, is almost certainly more difficult than form 7, which was employed on day 2, the improvement in ability of day 3 over day 2 is probably somewhat greater than is shown by the averages of our table and, yet, somewhat less than the improvement of day 2 over day 1.

Considering the separate tests, four of these (1, directions; 2, arithmetical problems; 3, practical judgment; and 6, number series completion) show continuously increasing averages throughout the experiment. Progress in the other four tests (4, synonym-antonym; 5, disarranged sentences; 7, analogies; and 8, information) is less regular. But only in test 8 was the highest average score obtained on day 1.

The greatest effects of practice are shown by tests 2 and 6. In the case of test 4 there is no appreciable improvement and in test 8 there is clearer evidence for a falling off in efficiency than for improvement.

Table III. presents the average number of points, problems or questions toward which the subjects gave a response which was either correct or incorrect, and also the relative achievement by this attempts criterion. According to the records of all the tests combined, Table III., like Table II., shows a considerable improvement for day 2 over day 1 and a less marked improvement for day 3 over day 2. As in the case of scores, variability decreases roughly in inverse proportion to improvement.

Five of the tests (2, arithmetical problems; 3, practical judgment; 4, synonym-antonym; 6, number series completion and 8, information) show an increasing number of points attempted throughout the experiment. Test 1, directions, shows neither positive nor negative practice effects. Positive effects would have been impossible in this case, due to the fact that the maximum number of attempts was made in the first day's performance. In the case of test 5, disarranged sentences, more attempts were made on day 2 than on day 3, but this difference is hardly significant. In test 7, analogies, there is a considerable superiority for day 2 over day 3. In no test, however, was the greatest number of attempts made on day 1.

Considering the evidence presented by all of our averages, then, the main trend is toward an improvement for day 2 over day 1, and somewhat less improvement for day 3 over day 2.

III. Consistency of Improvement among Individual Subjects

The inspection of Tables II. and III. demonstrated that, while in general there is a continuous improvement throughout the three days of the experiment, certain individual tests do not agree with our findings as a whole. Tables IV. and V. give the numbers of subjects, who, according to *scores* and *attempts*, improved continuously, lost continuously, or varied in achievement in some less regular fashion.

Table IV

Scores

3 means that this day's performance was lower than that of the other two days.

2 means that this day's performance was lower than that of one other day and higher than that of the other.

1 means that this day's performance was higher than that of the other two days.

Day 1	Day 2	Day 3	Number of Individuals Giving Days This Rank Order
3	2	1	14
3	1	2	17
2	3	1	4
2	1	3	1
1	2	3	1
1	3	2	2
Total......	....	....	39

Table V

Attempts

3, *2*, and *1* have same meaning as in Table IV.

Day 1	Day 2	Day 3	Number of Individuals Giving Days This Rank Order
3	2	1	17
3	1	2	9
2	3	1	3
2	1	3	1
1	3	2	1
2.5	1	2.5	2
1.5	3	1.5	1
3	1.5	1.5	5
Total......	....	....	39

Considering *scores* (Table IV.), the order 3–2–1, which stands for continuous improvement, is represented by fewer individuals than the order 3–1–2 which stands for initial improvement followed by a loss. There is only one subject, however, who shows a continuous loss throughout the experiment, and a decided majority (31 out of 39) of the subjects do better on both day 2 and day 3 than on day 1. Variations from the condition of continuous improvement are, therefore, due largely to the likelihood of day 2 showing a better performance than day 3. This is in accordance with the conclusion, based upon the averages of Table II., to the effect that the difference between day 1 and day 2 is probably greater than that between day 2 and day 3.

Considering *attempts* (Table V.), we find similar facts. The condition of continuous improvement here, however, is clearly most frequent. No subject shows a continuous loss, and again 31 out of 39 do better on days 2 and 3 than on day 1. These data further verify the generalization that practice in taking these tests is associated with improvement in achievement, especially of the second testing over the first.

IV. Effects of Practice on Functioning of Possible Critical Scores

The facts of improvement due to practice have so far been described in terms of changes in average improvement and in terms of the number of individuals whose achievement follows or varies from the general tendency of the group. For practical purposes it will be well to consider the specific effects which changes due to practice have upon the predictive value of the alpha examination. If, for example, *critical scores* for exclusion or selection were set on the basis of the group's performance on day 1, it would be of practical significance to note the effects of given amounts of practice upon the likelihood of such scores retaining their original exclusive or selective efficiency.

If, in terms of *scores*, that critical score had been selected upon the basis of the first day's records which would have separated the lower 9[1] from the upper 30 individuals, the

[1] The hypothetical critical scores which we shall discuss have been set to exclude the lowest 9 subjects or to select the 9 highest, except where there was a tie at the 9th

application of this critical score to the records of days 2 and 3 would have resulted in the exclusion on both days of only one individual, he being the same on both occasions. If a lower, or *exclusive*, critical score had been set on the basis of the records of day 2 so that the lowest 9 subjects on that day would have been excluded, the application of that score to the records of day 3 would have resulted in the exclusion of 7 subjects, 5 of whom belonged to the lowest 9 of day 2.

Similar results would have been obtained had lower critical scores been based on *attempts*. A score which would have excluded the lowest 9 on the first day would, if it had been applied to the second day's records, have excluded 2 individuals, but 1 of whom belonged to the originally excluded 9. Applied to the records of day 3, this critical score would have excluded only 1 individual.

Returning again to a consideration of *scores*, if an upper critical score had been set which would have selected the highest 9 subjects on day 1, that same score would have selected 16 subjects on day 2 and 18 subjects on day 3. If a critical score had been set which would have selected the highest 9 subjects on day 2, that score would have selected 11 subjects on day 3.

If *attempts* had been taken as the criterion and a critical score set which would have selected the highest 10 subjects on day 1, that score would have selected 19 subjects on day 2 and 24 subjects on day 3. If a critical score had been set which would have selected the highest 10 subjects on day 2, that score would have selected 13 subjects on day 3.

We may conclude from these facts that the changes in achievement due to even a small amount of practice, *i.e.*, that represented by taking the examination once, may largely invalidate the functional significance of a critical score based upon the records of a group having had little or no previous experience with the alpha examination. Since our evidence points to an early negative acceleration in the progress due

rank, in which case the critical score was set to exclude the lowest 10 or select the highest 10. These points have been chosen arbitrarily, of course, but they fall between the twenty-third and twenty-sixth percentile in the case of the lower one and between the seventy-fourth and seventy-fifth percentile in the case of the upper one.

to practice,[1] critical scores set after a group has had even a small amount of practice will probably maintain their significance fairly consistently for groups having had as much or more practice. This generalization is supported by the results of our application of critical scores based on second-day records to third-day results, and it lends support to the suggestion of Dunlap and Snyder[2] that if tests susceptible to practice are used, all candidates should have old forms of the test in advance. How much fore-exercise the candidates have is probably not so important as the fact that they should have some fore-exercise.

V. Effects of Practice upon Correlations Between Test Performances and Scholarship

The 39 subjects of our experiment were rated according to their scholarship records for five quarters, or one and two thirds college years. In Table VI. there are presented the values of

$$\rho\left(\rho = 1 - \frac{6\Sigma D^2}{n(n^2 - 1)}\right)$$

between the examinations of different days, and combinations of days, and scholarship. Where records of different days are combined, this was done in terms of ranks so as not to give extra weight to the performance on days most benefited by practice.

Table VI

Scores

	ρ	P.E.
Day 1 and Scholarship	.56	.07
Day 2 and Scholarship	.65	.06
Day 3 and Scholarship	.62	.07
Days 1 and 2, and Scholarship	.60	.07
Days 1 and 3, and Scholarship	.63	.07
Days 2 and 3, and Scholarship	.66	.06
Days 1, 2, and 3, and Scholarship	.64	.07
Attempts		
Day 1 and Scholarship	.41	.09
Day 2 and Scholarship	.32	.10

[1] The sharpness of this negative acceleration is somewhat doubtful according to our results, because of the impossibility of accurately determining the relative difficulties of forms 7 and 9, the forms used on the second and third days respectively.

[2] *Op. cit.*, p. 402.

Day 3 and Scholarship	.37	.09
Days 1 and 2, and Scholarship	.37	.09
Days 1 and 3, and Scholarship	.37	.09
Days 2 and 3, and Scholarship	.35	.10
Days 1, 2, and 3, and Scholarship	.37	.09

A consideration of the values in this table brings out the following facts:

1. There are noticeably higher correlations between *scores* and scholarship than between *attempts* and scholarship.

2. The differences between the coefficients secured on the different days of the testing are practically negligible. There is just a slight chance that the second or third day's examination, so far as *score* is concerned, has higher predictive value than that of day 1.

3. No significant increase in predictive value occurs when the test records of two or three days are combined.

In the previous section we considered the effects which one or two practice sittings would have upon the efficiency of a critical point set at some absolute score. It is also interesting to consider the relative eliminative or selective efficiency of certain critical ranks. Tables VII., VIII., IX. and X. present data bearing upon this point. Due to the comparatively low correlations between *attempts* and scholarship, these data are concerned only with scores.

Table VII

Students ranking from 33 to 39 inclusive in scholarship have grades which render them liable to elimination from college. Students having these ranks (33 to 39) in the tests of the different days and combinations of days have the following ranks in scholarship.

Days	Scholarship Ranks of Students Ranking from 33 to 39 in Tests
1	13.5, 28, 31, *33*, *37*, *38*
2	15, 20, 31, *33*, *34*, *38*, *39*
3	13.5, 28, 31, *33*, *37*, *38*, *39*
1 and 2	13.5, 28, 31, *33*, *37*, *38*, *39*
1 and 3	13.5, 28, 31, *33*, *37*, *38*, *39*
2 and 3	20, 28, 31, *33*, *37*, *38*, *39*
1, 2, and 3	13.5, 28, 31, *33*, *37*, *38*, *39*

Table VIII

Students ranking from 1 to 6 inclusive in scholarship have grades which render them eligible for *honors*. Students having these ranks (1 to 6) in the tests of the different days and combinations of days have the following ranks in scholarship.

Days	Scholarship Ranks of Students Ranking from 1 to 6 in Tests
1	*1*, *5.5*, 7, 8, 13.5, 22
2	*4*, *5.5*, *5.5*, 7, 16, 23.5
3	*3*, *4*, *5.5*, 7, 8, 23.5
1 and 2	*1*, *5.5*, 7, 8, 22, 23.5
1 and 3	*1*, *3*, *5.5*, 7, 8, 23.5
2 and 3	*4*, *5.5*, *5.5*, 7, 8, 23.5
1, 2, and 3	*1*, *5.5*, 7, 8, 16, 23.5

Table IX

Students ranking from 33 to 39 inclusive in scholarship have grades which render them liable to elimination from college. Students of this group are ranked as follows by the tests of the different days, and by combinations of days.

Days	Test Ranks of Students Ranking from 33 to 39 in Scholarship
1	17, 21, 21, 29.5, *34*, *36*, *39*
2	24, 31.5, 31.5, *33.5*, *35.5*, *37.5*, *39*
3	17.5, 24, 32, *33*, *35.5*, *37.5*, *39*
1 and 2	22, 23, 29.5, *33*, *35*, *38*, *39*
1 and 3	19, 22.5, 26, *33*, *35*, *36*, *39*
2 and 3	24.5, 26, 30, *33.5*, *35.5*, *38*, *39*
1, 2, and 3	22, 25, 27.5, *33*, *35*, *38*, *39*

Table X

Students ranking from 1 to 6 inclusive in scholarship have grades which render them eligible for *honors*. Students of this group are ranked as follows by the tests of the different days and combinations of days.

Days	Test Ranks of Students Ranking from 1 to 6 in Scholarship
1	*1*, *4*, 8, 14, 15, 17
2	*2*, *4*, *5*, 7.5, 12.5, 14.5
3	*1*, *5*, *6*, 7, 8, 10
1 and 2	*1.5*, *4*, 8.5, 8.5, 10, 16
1 and 3	*1*, *5*, *6*, 10, 12.5, 12.5
2 and 3	*1*, *4*, *6*, 7, 9, 10.5
1, 2, and 3	*1*, *5*, 7, 9, 10, 12

According to the present interpretation of scholarship records by the administration of this college, students whose grades are as low or lower than those of the lowest 7 members of our group are eligible for dismissal. For practical purposes one might, in light of this fact, set a critical rank in test performance which would separate the lowest 7 individuals from the group as a whole.[1] Table VII. shows what are the ranks in scholarship obtained by the 7 individuals who are lowest according to the examinations of the different days and combinations of days. Only 6 individuals are considered in the case of day 1, because of a tie at the critical rank.

None of the differences between these test measures is large enough to justify any confident statement of superiority for particular days or combinations of days. There is a slight chance that day 1 is not quite as good a selective agency as the other days, which apparently are of equal value.

Those students of our group ranking from 1 to 6 inclusive in scholarship have grades which render them eligible for *honors*. If we were interested in selecting honor students, we might set an upper critical rank which would select the best 6 members of the group. Table VIII. shows what scholarship ranks are held by those individuals who are included in the best 6 by the tests of different days and combinations of days. There is here a slight chance that day 1, days 1 and 2 combined, and days, 1, 2, and 3 combined are less efficient than day 2, day 3, days 1 and 3 combined, and days 2 and 3 combined, but these differences again may easily be due to chance.

Tables IX. and X. show the test ranks according to the different days and combinations thereof of the lowest 7 and the highest 6 individuals according to scholarship. In other words, they show which individuals would have been eliminated who, according to scholarship, should have been eliminated, and which individuals would have been marked for *honors* who should have been so marked. These results, like those of Tables VII. and VIII., give little ground for choice among the different test measures.

[1] It is theoretically possible that better results might be achieved by dividing the group at a point slightly above or below this one.

VI. Causes of Improvement

An effort to analyze the different factors which may have influence in accounting for the variability with practice reveals a high degree of complication. Various intellectual and emotional changes, and changes in method must be considered, along with such an external influence as the increase in time afforded by the fact that the subjects may neglect most of the instructions given by the examiner on the second and third days.

This extra time available during the second and third tests may be an important factor in determining the cause of an *apparent* practice effect. Excepting test 1, oral directions, where the instructions are given for the different items consecutively, the total time for all of the tests (2–8) is 20 minutes. The instructions, as read to this group, consumed a total of 4′ 38″ or 23.1 per cent. of the time allowed for work. When given according to the manual, the subject looks at the directions at the beginning of each test as the examiner reads them aloud. At the time of the first trial the subject is engrossed in the instructions; on the later trials, he is likely to feel so familiar with them, that he has the reading time free to dispose of otherwise. This added 23 per cent. of his gross time may be used in working ahead mentally on the actual test before him, so that he is well on his way when the signal to begin is given. When this time was not used as additional work time, it was often taken as a relaxation period. In introspective reports on the effects of practice, about one third of those reporting explicitly on this added time, spoke of taking it wittingly as a recuperative period. About two thirds used it for work.

This additional time might well be taken into account and controlled when similar tests are repeated. The plan was considered of reading the directions, on the later days, before the subject was permitted to turn the page to the approaching test. That would have meant, however, the substitution of indicated rest intervals amounting to nearly one fourth the time spent on the tests on the first day. An alternative of omitting, after the first examination, the reading of the directions would have meant reducing the gross time for the seven tests from about twenty-five minutes to a little more than twenty, allowing for time consumed in turning pages. This seemed inadvisable, as it introduced a change in method, the effect of which could not be estimated.

An important practice effect, as *practice*, seemed to be due to the development of method during and after the first day's work and its carrying over to later work. Many students spoke of a 'systematization' that was advantageous in some of the tests, particularly in test 2, arithmetical problems, and test 6, number series completion, which show continuously increasing average scores from day to day. The methods employed, in the case of test 6, were the same in succeeding problems on different days, so that each day's work pushed the frontier of progress a little further down the page. It was to this frontier that a number of students devoted themselves during the reading of the directions on days 2 and 3. This increase in effectiveness in working due to the development of method probably characterizes every test employed in alpha, although in some cases as in test 8, information, the method may largely be concerned with such activities as the eye-hand coördinations involved in crossing out words.

Another important factor was the change, on days 2 and 3, in the attitude toward the tests. This worked in two ways. The first effects, the disappearance of nervousness and anxiety, and the establishment of confidence were all conducive to an increase in score. The second effects were overconfidence, loss of interest and explicit boredom which in a number of cases were related to a decrease in the scores for the third day as compared to day 2. The general appearance of the class indicated this lack of interest in the third test.

Some transfer effect was of an intellectual sort, notably the increase in ability, in test 5, disarranged sentences, to take in at a glance the meaning of the sentence. A part of this perceptual factor may be accounted for as a method. The memory of earlier solutions in the arithmetical problems also carried over.

The high percentage of improvement in test 6, number series completion, second only to that in test 2, arithmetical problems, is doubtless due to the partial identity of the test materials in any two forms of alpha. In taking test 6 (form 7) on the second day the subjects found, in five of the twenty items, numbers of the same sequential series, *e.g.*, 2, 3, 4, 5, 6, 7, but beginning in a different point in the series, *e.g.*, 3, 4, 5, 6, 7, 8. In the remaining 15 items he found different *numbers*, but the same *methods* of completion, *e.g.*, 8, 1, 6, 1, 4, 1, and 9, 1, 7, 1, 5, 1. On the third day, all of the 20 items were identical with items of the two previous days and presented no new material.

This fact, of course, indicates a difficulty in equating different forms of such tests, due to the types of material used.

There was also some indication of added practice due to discussions between the examinations on various days. This discussion concerned itself with methods of solving arithmetical problems, and with verifications of answers in the information test, as, for instance, the apparently controversial point as to the number of a Zulu's legs. Here the dictionary and the encyclopedia contributed a basis for estimating the biological classification of the Hottentot and the Papuan.

The fact that absence of improvement, or even loss, sometimes occurs is probably due to the facts that the different forms of the alpha examination were equated as wholes and not part with part, and that in the case of some parts of the test the subject either has or has not the information necessary for a correct response; *i.e.*, some parts of the test are informational rather than problem solving in nature.

VII. Summary

1. Practice effects in the alpha examination administered in different forms to college students on three successive days show a marked improvement for day 2 over day 1 and probably a less marked improvement for day 3 over day 2. This result is an average tendency rather than a universal characteristic of all parts of the test and of the records of all subjects.

2. The amount of practice represented by one experience with the alpha examination largely invalidates critical scores

of a given selective or exclusive significance set upon the basis of a first performance with this test. Critical scores set upon the basis of a second performance are more likely to maintain their functional significance in the face of further practice.

3. Performance according to *scores* in all cases correlates more highly with scholarship than performance according to *attempts*.

4. There are no very important differences between the test-scholarship correlations for the different days or combinations of days. The *scores* of day 1 give the least accurate prediction of scholarship, but this inferiority may easily be due to chance.

5. Improvement is due to an increase in available time on succeeding tests due to the possible disregard of instructions; to an increase in motor facility; to development of methods; to shift in emotional attitude; to a transfer in perceptual factors; to identity of materials in the several forms, and to outside discussion of various points between examinations.

THE TASTE SENSITIVITY OF AN ANOSMIC SUBJECT

BY A. R. GILLILAND

Dartmouth College

When one of the special senses is missing or lost is there compensation in closely related senses? This question has received much theoretical consideration but little exact experimental study. A recent class demonstration of the importance of smell in so-called taste sensations brought out the fact that one member of the class had no sense of smell. A freshly cut piece of onion held to his nose was undetected so long as the eyes were closed. Experiments in chemistry were sometimes very difficult but the subject was generally able to make use of other means than smell for detecting chemical differences. In fact he did his work so well that his instructor did not know that the student was anosmic. The anosmic condition has been present from birth.

Such a subject presents an opportunity for a study of two interesting problems: (1) Is there compensation in two such closely related senses as taste and smell when one of these is lacking? If there be compensation of function, at least other than that resulting from increased exercise of the remaining parts or functions, it would seem likely to be present in two such closely related senses as these. (2) Whether or not there be compensation, is there any greater relative capacity to detect complex tastes by the sense of taste alone for such a subject? That is, will the greater reliance on taste by the anosmic make him able to detect different foods better than the normal subject when not using the sense of smell?

Solutions of cane sugar, quinine, tartaric acid, and sodium chloride, approximately four times the strength of the limens reported by Baily and Nicols[1] were prepared, for preliminary

[1] 'The Delicacy of the Sense of Taste,' *Nature*, XXXVII., 1887–88, 557.

experiments for determining the limen for each of the four primary tastes. The solutions were placed in tumblers on a table. The subject sat with his back to the table so that he did not know what substance was being given to him. The order of presentation was unknown to the subject. The experimenter gave the subject a teaspoonfull of the substance and the subject reported the taste. Five trials for each solution were made with as many added trials of ordinary tap water. This made forty trials in each experiment, twenty for water and five for each of the four solutions. The liquids were all of ordinary room temperature, about 30 degrees Centigrade. Four subjects took the experiment, the anosmic and three control subjects.

The original solutions were found to be entirely too strong. In all, twenty trials were made, five for each subject, before the best limen for each solution was determined. The results for the last six experiments are presented in Tables I. and II. The first table gives the responses and the scores for three subjects with the following strengths of solutions.

Sour (tartaric acid)	1 part in 25,000 parts of water
Bitter (quinine)	1 part in 300,000 parts of water
Salt (sodium chloride)	1 part in 6,250 parts of water
Sweet (cane sugar)	1 part in 320 parts of water

The second table gives the results for three subjects with salt 1 part in 9,375 parts of water instead of the strength given above. The other solutions were the same strength as before. This proved to be very near the limen for these three subjects and probably nearer the limen for the other subject who did not try this series of tests.

The scores given are for correct responses. If the subject was certain that he tasted one of the primary tastes he reported which one it was. If he did not taste anything in the liquid he reported the taste as water. Very often the subject was not certain. He sometimes reported that it was probably one or another of the substances or that it was possibly one of them. Such reports were not credited even if correct. Only unequivocally correct responses were scored.

Table I

Stimulus	Strength	Subject* *W*	Score	Subject* *P*	Score	Subject* *G*	Score
Acid (Tartaric)	$\frac{40}{1000000}$	bitter		water		wate	
		bitter		bitter		acid	
		acid	3	water	0	acid	2
		acid		bitter		acid?	
		acid		bitter		bitter	
Bitter (Quinine)	$\frac{3\frac{1}{3}}{1000000}$	acid		salt		bitter	
		water		salt		???	
		acid	0	bitter	1	bitter	1
		acid?		acid		water	
		sweet		acid		water	
Salt (NaCl)	$\frac{160}{1000000}$	water?		salt		salt	
		salt		salt		salt	
		salt	4	acid	4	salt	5
		salt		salt		salt	
		salt		salt		salt	
Sweet (Sugar)	$\frac{3125}{1000000}$	sweet		sweet		water	
		sweet		water		water	
		sweet	4	sweet	4	sweet	1
		bitter		sweet		water	
		sweet		sweet		sweet?	
Water		water		water		acid°	
		bitter		water		water	
		water		water		water	
		water		water		water	
		water		water		something	
		water		water		water	
		acid?		salt		bitter?	
		water		water		sweet	
		bitter		water		salt?	
		water		water		water	
		sweet?	8	water	17	water	14
		water		water		bitter	
		bitter?		water		water	
		acid		bitter		water	
		bitter		acid		water	
		water?		water		water	
		sweet		water		water	
		acid		water		water	
		water?		water		water	
		salt		water		water	
		Total Score	19		26		23

TABLE II

Stimulus	Strength	Subject* *W*	Score	Subject* *M*	Score	Subject* *G*	Score
ACID	$\frac{40}{1000000}$	sweet?		water		sweet	
(Tartaric)		bitter?		bitter		water	
		acid	1	acid?	1	bitter?	0
		water		acid		water	
		sweet		bitter		bitter	
BITTER	$\frac{3\frac{1}{3}}{1000000}$	water		bitter		water	
(Quinine)		water		bitter		bitter	
		bitter	1	bitter	4	water	1
		water		acid		water	
		acid		bitter		water	
SALT	$\frac{107}{1000000}$	bitter		salt		water	
(NaCl)		acid		water		salt	
		salt	2	salt?	2	water	2
		bitter		salt		water	
		salt		salt?		salt	
SWEET	$\frac{3125}{1000000}$	salt?		water		sweet	
(Sugar)		water		sweet?		water	
		sweet	1	water	0	salt	3
		water		a taste		sweet	
		salt		water		sweet	
WATER		bitter		water		acid?	
		water		water		water	
		water		water		water	
		acid		water		water	
		salt		water		acid	
		acid		water		bitter?	
		sweet		water		bitter	
		water		salt		water	
		water		water		bitter?	
		sweet?	9	water	18	water	14
		water		water		sweet	
		water		acid		water	
		water		water		water	
		bitter		water		water	
		water		water		water	
		sweet		water		water	
		bitter		water		water	
		acid		water		water	
		water		water		water	
		acid		water		water	
		Total Score	14		25		20

Very often water was reported as having a taste. Distilled water was used at first but it seemed to prove more objectionable than tap water. With either it was impossible to eliminate this difficulty except with one subject who almost invariably reported water correctly. The anosmic subject reported a taste or a possible taste incorrectly for water more than half the time. For this reason water was used as a control as many times as all the other substances combined. The total possible score, one point for each correct response, was therefore 40.

The problem of fatigue of the end organs during the experiment was not so serious as might have been expected. When the limen of sensitivity was approached the different substances were so weak that their after effects were very small. The mouth was thoroughly rinsed several times during the experiment, at least once after each report of a definite taste. The frequent use of water as a control also helped to keep the mouth thoroughly cleansed. The fact that the latter part of the experiment contained as many correct responses as the first part is evidence that fatigue was not an important distracting element in the experiment.

The results show that the anosmic subject, *W*, was less able to detect the four primary tastes than any of the normal subjects. In the three records of Table I. the anosmic subject's score was 19 while one of the subjects, *G*, made 23 and the other, subject *P*, made 26. In Table II. subject *W*, the anosmic, made a score of 14; subject *G* made 20; and subject *M* made 25.

There is not a complete agreement in the scores of the same subject in the two tables. At least one cause besides accidental variation tended to produce this difference. In the first set of tests the salt solution was clearly above the limen. This not only improved the score for salt but left less chance for error with the three remaining primary tastes. The subject's procedure when the solution was very weak, was at least partly one of elimination:—first is or is not the substance one of the primary tastes? Second, if so which one is it? In the first set of experiments both the first and second judgments were rendered easier because salt was readily eliminated as a possible taste. It was always clearly present or absent.

When the solutions were very weak it became largely a matter of judgment from a large number of external and internal factors. It was surprising how eagerly the subjects grasped for any possible clue for trying to detect the taste. It was necessary that the liquids all be of the same temperature. The experimenter had to be especially careful not to strike the tumblers with the spoon as the direction of the sound was immediately used by the subject as a cue in trying to differentiate the tastes. The order of presenting the different liquids also had to be very carefully worked out so as to allow the least possible chance for the subject's guessing the order.

We have considered the limen as the point at which the judgments were just perceptibly better than might be expected by the theory of chance. We might have chosen the point at which no errors in report were made or some other per cent. of success.

Since the different subjects did not have exactly the same limen for each of the primary tastes the strengths given in Table II. is approximately an average of the limens for the four subjects. Subject *M* probably could have detected bitter in a weaker solution. Subjects *P* and *G* might have reduced the limen for sweet but their two records do not show enough consistency to warrant a much greater reduction. Our final results do not differ greatly from the limens given by Baily and Nicols.[1]

Our second problem was to determine whether the anosmic subject could distinguish complex tastes—ordinary foods—better than a normal subject using the sense of taste alone. Seven raw foods were selected for the experiment—cabbage, apple, Irish potato, sweet potato, onion, carrot, and turnip. The anosmic and ten other students acted as subjects in this part of the experiment. The subjects sat with their back to the table containing the foods. Each food was presented

[1] *Op. cit.*

five times to each subject. Just as the particle of food was being presented to the subject he closed his nostrils with his fingers and locked his breath by filling the nasal cavity with air. In this condition practically no odor could reach the olfactory area so long as the subject held his breath. The subject was allowed to chew the substance as he would do in eating but was cautioned to hold his breath until his report was given.

The same method of scoring was used as described in the first experiment. The total possible score was 35. Table III. gives the responses and scores for the eleven subjects.

In order to be sure that unfamiliarity with taste of the seven foods since some of them are seldom tasted except when cooked a supplementary experiment was conducted with the anosmic and three other subjects. Five common raw foods were selected—orange, grapefruit, pineapple, tangerine, and tomato. Five trials were used for each. Out of a possible score of 25 the anosmic subject made 18, another 18, and the other two 21 each. The relative scores for each subject was practically the same as in the preceding experiment.

These results show that the anosmic subject was not better than the average normal subject; in fact not quite as good as the average for eleven cases. The much greater experience of the anosmic did not result in his ability to distinguish different kinds of complex tastes better than normal subjects using the sense of taste alone.

Conclusions

So far as we may generalize from the results of a single case there is no greater sensitivity to the primary tastes with an anosmic subject to compensate for the lack of the sense of smell.

The anosmic subject had no greater ability in detecting the complex tastes than the normal subject when using the sense of taste alone.

These results tend to disprove rather than prove a theory of the compensation of function.

TABLE III—*Continued*

Stimulus	Response											
Subject	W (Anosmic)	P	C	M	J	S	Cs	Me	G	Py	H	Sn
CABBAGE	turnip?	carrot	apple	onion	onion	cabbage	apple	cabbage	cabbage	cabbage	cabbage	onion
	carrot	carrot	cabbage	cabbage	cabbage	cabbage	cabbage	???	cabbage	carrot	cabbage	cabbage
	cabbage	cabbage	cabbage	cabbage	cabbage	turnip	turnip	I. potato	cabbage	cabbage	cabbage	cabbage
	cabbage	carrot	turnip	cabbage	S. potato	cabbage	cabbage	???	onion	cabbage	cabbage	cabbage
	turnip	cabbage	onion	S. potato	S. potato	turnip	cabbage	cabbage	cabbage	cabbage	cabbage	cabbage
APPLE	apple	apple	apple	onion	apple	apple	apple	apple	apple	apple	apple	apple
	apple	potato	turnip	apple	apple	apple	apple	apple	apple	apple	apple	apple
	apple	apple	apple	apple	apple	apple	apple	apple	apple	apple	apple	apple
	apple	apple	apple	apple	apple	apple	apple	apple	apple	apple	apple	apple
	apple	turnip	apple	apple	apple	apple	apple	apple	apple	apple	apple	apple
IRISH POTATO	I. potato	onion	apple	onion	onion	S. potato	I. potato	I. potato	I. potato	I. potato	I. potato	I. potato
	I. potato	onion	I. potato	turnip	I. potato	I. potato	I. potato	I. potato	I. potato	I. potato	S. potato	I. potato
	I. potato	dirt	cabbage	I. potato	I. potato	cabbage	I. potato	???	I. potato	I. potato	I. potato	I. potato
	I. potato	???	cabbage	onion	I. potato	I. potato	I. potato	I. potato	I. potato	I. potato	S. potato	I. potato
	I. potato	???	turnip	I. potato	I. potato	S. potato	I. potato	I. potato	I. potato	S. potato	S. potato	I. potato
SWEET POTATO	turnip?	carrot	turnip	apple	I. potato	S. potato	turnip	S. potato	S. potato	turnip	turnip	cabbage
	S. potato	carrot	S. potato	carrot	S. potato	I. potato	carrot	S. potato	carrot	S. potato	S. potato	I. potato
	turnip	S. potato	S. potato	carrot	S. potato	S. potato	carrot	S. potato	S. potato	I. potato	turnip	S. potato
	turnip	cabbage	S. potato	carrot	carrot	I. potato	turnip	S. potato	S. potato	S. potato	turnip	I. potato
	turnip	???	apple	carrot	S. potato	S. potato	S. potato	S. potato	S. potato	S. potato	S. potato	S. potato

Stimulus / Subject	W (Anosmic)	P	C	M	J	S	Cs	Me	G	Py	H	Sn
	Response											
Onion	onion	onion	cabbage	turnip	turnip	onion	apple	onion	turnip	onion	onion	onion
	onion	onion	onion	onion	carrot	cabbage	turnip	onion	I. potato	onion	onion	onion
	onion	onion	onion	cabbage	turnip	cabbage	onion	onion	onion	apple	onion	onion
	onion	onion	apple	onion	onion	onion	onion	onion	onion	cabbage	onion	onion
	onion	onion	onion	onion	carrot	onion	apple	onion	onion	onion	apple	onion
Carrot	turnip	cabbage	turnip	carrot	turnip	turnip	carrot	S. potato	carrot	S. potato	cabbage	turnip
	turnip	I. potato	potato	cabbage	S. potato	carrot	S. potato	???	turnip	carrot	cabbage	carrot
	turnip	I. potato	cabbage	turnip	carrot	turnip	carrot	cabbage	carrot	S. potato	cabbage	carrot
	carrot	carrot	I. potato	I. potato	turnip	I. potato	carrot	carrot	cabbage	carrot	carrot	carrot
	cabbage	cabbage	turnip	turnip	turnip	turnip	carrot	cabbage	cabbage	turnip	cabbage	carrot
Turnip	carrot?	???	I. potato	cabbage	onion	I. potato	I. potato	turnip	turnip	S. potato	turnip	apple
	carrot	cabbage	carrot	I. potato	onion	turnip	I. potato	S. potato	onion	turnip	turnip	turnip
	turnip	I. potato	I. potato	carrot	turnip	carrot	S. potato	cabbage	cabbage	turnip	carrot	turnip
	cabbage	I. potato	I. potato	turnip	turnip	cabbage	turnip	???	turnip	turnip	carrot	turnip
	I. potato	I. potato	I. potato	S. potato	cabbage	turnip	carrot	???	onion	carrot	turnip	turnip
Score	19	12	13	14	18	19	21	23	25	24	25	29

Journal of

Experimental Psychology

VOL. IV, No. 5 OCTOBER, 1921

THE RELATIVE EFFICIENCIES OF DISTRIBUTED AND CONCENTRATED STUDY IN MEMORIZING

BY EDWARD S. ROBINSON

The University of Chicago

THE PROBLEM

It is a well-known fact, supported by the results of thirty or more researches,[1] that efficiency of learning is a function of the manner in which practice is distributed. The literature on this subject raises several more or less distinct questions: (1) What is the relative efficiency of concentrating all practice or study into one sitting as compared with dividing it into smaller units indulged in at varying intervals of time? (2) Granted that the total amount of practice is to be divided, what are the optimal units of practice? (3) What are the optimal intervals between given amounts of practice? The experiments to be reported in the present paper are concerned with the first of these questions and some of those factors which determine its answer for the memorizing of numbers.[2]

It has already been demonstrated that the relative efficiencies of concentrated and distributed study depend upon such factors as the nature of the material,[3] the length of the material,[4] and various features of the method of study.[5] In all

[1] See bibliography at end of paper.

[2] It is logically possible to fit the present experiments under the third as well as under the first of these questions. In fact, the results obtained throw some light also on the second question. This is due to the fact that these questions are all very closely related.

[3] Lyon, *J. Ed. Psychol.*, 1914, 5, pp. 1–9, 85–91, 155–163.

[4] Lyon, *op. cit.*

[5] See for example, Mould, Treadwell, and Washburn, *Amer. J. of Psychol.*, 1915, 26, pp. 286–288.

likelihood the relative efficiencies of concentration and distribution are somewhat influenced by many other factors, such as the degree of learning considered, methods of measuring efficiency (whether amount, accuracy or speed of recall is the criterion), the time which has elapsed since the final reading of the material, and the like. The effects of some of these factors are examined in the experiments which follow.

The Method

Two experiments were conducted. The memorized materials in both cases were lists of ten three-place numbers. The lists were made up by chance, except that:

1. In no number was 0 employed.
2. No digit within a number was immediately followed by a digit one more in value.
3. The same digit never occurred twice in the same number.
4. No digit in one number appeared in the immediately succeeding or following number.
5. The numbers 135, 357, 579, 246, and 468 were not employed.

In both experiments the numbers of any list were exposed for 2 seconds each with no appreciable interval between the covering of one number and the exposing of the succeeding number within a list. Between the immediately successive exposures or presentations of a complete list there was an interval of approximately 6 seconds.

In the first experiment each list of numbers was written on a long slip of paper which could be drawn through a double slit in a piece of white cardboard in such a way that only one number of the list came into view at a time. Timing was controlled with reference to a pendulum which swung on the wall behind the subject. In the second experiment the numbers were exposed by means of a hand-operated memory drum and timing was with reference to the clicks of a metronome.

Each experiment contained six conditions. In Experiment I., for example, a list of ten three-place numbers was given either 12 presentations at a sitting, or 6 presentations at one sitting and 6 more presentations 24 hours later. Recall for each study method was measured at 5 minutes, 20 minutes, or 24 hours after the completion of study. The conditions of the two experiments are displayed in Table I.

Table I

Experiment I.

Condition A.—12 presentations of list of numbers, interval of 5 minutes, recall.
Condition B.—12 presentations of list of numbers, interval of 20 minutes, recall.
Condition C.—12 presentations of list of numbers, interval of 24 hours, recall.
Condition D.—6 presentations of list of numbers, interval of 24 hours, 6 more presentations of same list, interval of 5 minutes, recall.
Condition E.—6 presentations of list of numbers, interval of 24 hours, 6 more presentations of same list, interval of 20 minutes, recall.
Condition F.—6 presentations of list of numbers, interval of 24 hours, 6 more presentations of same list, interval of 24 hours, recall.

Experiment II.

Condition A.—6 presentations of list of numbers, interval of 5 minutes, recall.
Condition B.—6 presentations of list of numbers, interval of 20 minutes, recall.
Condition C.—6 presentations of list of numbers, interval of 24 hours, recall.
Condition D.—3 presentations of list of numbers, interval of 24 hours, 3 more presentations of same list, interval of 5 minutes, recall.
Condition E.—3 presentations of list of numbers, interval of 24 hours, 3 more presentations of same list, interval of 20 minutes, recall.
Condition F.—3 presentations of list of numbers, interval of 24 hours, 3 more presentations of same list, interval of 24 hours, recall.

The activity of the subjects between study and recall was uncontrolled, except for the fact that the subjects were asked to refrain from rehearsing the numbers during that interval.

Recall was written. The experimenter handed the subject a record book, open at a blank page, and said:

"I want you to write here in proper order as many as you can of the numbers you studied. Work as quickly and accurately as you can. Tell me as soon as you are through writing."

Since each subject was asked for a recall on 12 or 24 different occasions, the directions were dropped as soon as it was clear that the subject realized what was expected of him. I have found in previous work of this character that the constant repetition of such simple directions is either annoying or highly amusing to the subject, and, for that reason, their omission probably introduces less error and inconstancy than their continuation.

The speed of recall was taken by means of a stop watch which was started as the blank page was placed before the subject and stopped when he indicated that he had reproduced as much of the material as he could.

Each subject employed was tested at least twice under all of the conditions of one experiment, and before he began the study of a new list of numbers he was always informed as to the nature of the condition before him.

Experiment I. employed 12 subjects and Experiment II. employed 6. The subjects in both cases were arranged into one or two groups of 6, and each subject in each group was put through the conditions in different order. The arrangement used is shown in Table II.

Table II.

Order in Which Different Subjects Were Tested Under the Different Experimental Conditions

Order	1st	2d	3d	4th	5th	6th
Subject 1..........	A	B	C	D	E	F
Subject 2..........	B	C	D	E	F	A
Subject 3..........	C	D	E	F	A	B
Subject 4..........	D	E	F	A	B	C
Subject 5..........	E	F	A	B	C	D
Subject 6..........	F	A	B	C	D	E

Each of the two groups of subjects in Experiment I. went through two cycles as represented in Table II. The one group in Experiment II. went through four cycles. This gave a total of 24 separate tests for every experimental condition.

Obviously the arrangement of conditions adopted was an attempt at least partially to equate the amounts of previous practice effective for the different conditions. A further precaution was taken in that the subjects were given three practice sittings where lists of ten three-place numbers were studied and recalled. It also was so arranged that each list of numbers used was studied by some subject under each of the experimental conditions. This should have compensated in some degree for possible inequalities in the difficulties of the lists.

At no sitting did a subject work with more than one list of numbers. A subject was never asked, for instance, to begin memorizing a new list of numbers at a sitting during which he had previously studied or recalled another list. It was my intention here to guard as carefully as possible against positive or negative transfer between the lists. Of course, where a subject works day after day with units of similar material there must be considerable transfer of which it is practically impossible to take account. It seems reasonable to suppose, however, that the amount of transfer between the different units of material will be greatly reduced if not more than one unit is dealt with at a sitting.

Not all of my subjects were tested at the same time of day. Such an arrangement might have been desirable, but it was hardly practicable. Nevertheless, each subject had a regular sitting time, and this was kept constant with the exception of a few variations of not over two or three hours.

The subjects in Experiment I. were undergraduate and graduate students in Yale University, none of whom had had extended experience of this sort. Five graduate students and one undergraduate in the University of Chicago served as subjects in Experiment II.

A number of Yale University students in the writer's class in experimental psychology acted as experimenters in Experiment I. Their work was carried out under detailed instructions and I have no reason to feel anything but confidence in its accuracy. The writer was the only experimenter in Experiment II.

The criteria of recall were three: (1) Total Digits—the number of digits recalled irrespective of their correctness; (2) Correct Digits—the number of digits recalled which were correct in both content and position (where a three-place number was removed by not more than one place from its correct position a penalty for one incorrect digit was applied); (3) Time—the total time taken for recall divided by the total number of digits recalled.

Results

Table III. presents the average performances under the conditions of Experiment I. Columns bearing the title *6 + 6* contain the average performances where the memorized material was given 6 presentations at one sitting and 6 presentations 24 hours later. The columns headed *12* contain the average performances where the memorized material was given all 12 presentations at one sitting. Table IV. presents the averages under the conditions of Experiment II. The notation is analogous to that of Table III. With two exceptions each average in these tables represents 24 individual measurements. One of these exceptions represents 22 measurements and the other 23.

Table III

Experiment I. (Means of All Individuals and Both Cycles)

	Recall after 5 Minutes		Recall after 20 Minutes		Recall after 24 Hours	
	6 + 6	*12*	*6 + 6*	*12*	*6 + 6*	*12*
Total digits	25.4	24.6	24.8	25.0	21.6	20.3
Correct digits	23.1	20.8	21.3	20.2	15.4	9.6
Time per digit (seconds)	1.85	2.55	2.15	2.25	2.8	3.4

Table IV

Experiment II. (Means of All Individuals and All Cycles)

	Recall after 5 Minutes		Recall after 20 Minutes		Recall after 24 Hours	
	3 + 3	*6*	*3 + 3*	*6*	*3 + 3*	*6*
Total digits	19.3	20.4	19.1	19.7	14.5	14.0
Correct digits	14.6	16.9	13.8	14.6	8.8	6.8
Time per digit (seconds)	2.6	2.2	2.2	2.8	3.92	3.95

According to eight out of the nine pairs of values in Table III. the distributed form of study is superior to the more concentrated form. There is little evidence in these unevaluated averages concerning the effects which the progress of forgetting or the criterion of measurement employed may have upon this superiority.

The averages from Experiment II. (Table IV.) show no such clear advantage for the distributed study. In fact, five out of nine pairs favor concentrated study. Evidently the effects of dividing a series of presentations into halves 24 hours apart are dependent upon the total number of presentations employed. The dividing into halves of 12 presentations is advantageous; the dividing into halves of 6 presentations is, if anything, disadvantageous.

While, in Experiment I., so far as mere averages are concerned, there seems to be no significant change in the relative results of the two methods of study with the progress of forgetting, the averages from Experiment II. tell a different story. At 5 minutes and 20 minutes after the final presentation, the concentrated method shows the greater efficiency, but at the end of 24 hours this condition is reversed and all three[1] criteria favor distribution.

TABLE V

RESULTS OF EXPERIMENT I.—RECALL AFTER 5 MINUTES

First Cycle

Subjects	Total Digits		Correct Digits		Time per Digit (Seconds)	
	6 + 6	*12*	*6 + 6*	*12*	*6 + 6*	*12*
1	30	24	26	21	1.4	1.0
2	15	18	15	18	3.0	1.9
3	30	30	28	22	1.8	2.7
4	30	30	28	19	1.5	2.8
5	30	27	27	23	1.4	2.5
6	18	9	17	7	2.0	3.9
7	23	16	16	10	1.5	3.7
8	30	30	26	26	3.3	5.0
9	30	30	30	30	1.0	1.0
10	27	21	25	9	1.8	2.8
11	18	9	13	7	1.4	5.5
12	24	30	18	20	2.5	2.0
Mean	25.4	22.8	22.4	17.7	1.9	2.9

No attempt has been made to evaluate the means in Tables III. and IV. by the use of the conventional methods of calculating $P.E._m$ or $S.E._m$. It is quite evident that the assumptions underlying these methods render them inapplicable to the data of these experiments. Table V., VI., and

[1] In the case of *time of recall* there is probably no significant difference between the two methods.

Second Cycle

1	20	24	17	22	2.7	1.0
2	15	21	15	21	1.2	1.7
3	30	30	29	25	2.3	2.0
4	30	30	30	24	1.6	2.3
5	30	27	30	23	0.7	2.5
6	15	12	12	10	2.3	3.7
7	19	24	17	24	1.8	1.4
8	30	30	27	28	2.3	1.3
9	30	30	30	30	0.7	2.7
10	30	30	30	26	2.0	2.7
11	26	28	25	26	1.5	1.9
12	30	30	24	27	2.0	3.5
Mean	25.4	26.3	23.8	23.8	1.8	2.2
No. times superior	8	6	13	8	16	7
No. ties	10		3		1	

Total no. times *6 + 6* is superior = *37*
Total no. times *12* is superior = *21*
Total no. ties " = *14*

Table VI

Results of Experiment I.—Recall after 20 Minutes

First Cycle

Subjects	Total Digits		Correct Digits		Time per Digit (Seconds)	
	6 + 6	*12*	*6 + 6*	*12*	*6 + 6*	*12*
1	15	30	13	15	2.6	1.1
2	15	15	15	13	1.3	2.2
3	30	30	26	26	1.5	2.2
4	30	30	30	16	2.3	3.1
5	30	30	22	25	1.4	1.7
6	12	18	6	9	1.5	2.2
7	19	24	14	19	3.1	1.9
8	27	27	21	17	3.3	4.1
9	30	30	30	30	1.8	1.3
10	30	24	26	16	1.3	1.2
11	18	16	15	10	1.6	1.9
12	27	30	17	24	1.5	4.0
Mean	23.6	25.3	19.6	18.3	1.9	2.2

VII., Experiment I., and Tables VIII., IX., and X., however, include the detailed results of those of Experiment II. An examination of these tables will throw light upon the validity of the generalizations made upon more general bases and, perhaps, reveal some important facts not apparent in Tables III. and IV.

Each table contains 72 pairs of values, the members of which can be compared in estimating the relative efficiencies of the two methods of study. Pooling all of the comparisons for Experiment I., Tables V., VI., and VII., it appears that:

Distributed study is superior 114 times
Concentrated study is superior 63 times
Neither is superior 39 times

Second Cycle

1	30	24	28	23	1.0	1.6
2	15	18	15	16	6.6	1.9
3	30	30	27	22	2.5	2.4
4	30	30	30	26	1.6	2.5
5	30	30	30	25	0.8	1.7
6	15	18	15	16	2.0	2.3
7	25	12	21	10	1.2	2.2
8	30	16	21	14	2.3	5.0
9	30	30	30	30	1.8	1.1
10	27	30	19	26	3.7	1.3
11	19	28	16	27	2.3	0.9
12	30	30	24	29	3.0	4.7
Mean	25.9	24.7	23.0	22.0	2.4	2.3
No. times superior	5	8	11	10	15	9
No. ties	11		3		0	

Total no. times $6 + 6$ is superior $= 31$
Total no. times *12* is superior $= 27$
Total no. ties $= 14$

Table VII

Results of Experiment I.—Recall after 24 Hours

First Cycle

Subjects	Total Digits		Correct Digits		Time per Digit (Seconds)	
	$6 + 6$	12	$6 + 6$	12	$6 + 6$	12
1	15	24	9	9	2.0	1.6
2	15	15	6	7	6.9	9.0
3	30	30	18	14	1.8	2.8
4	30	30	30	14	2.8	3.3
5	30	30	17	15	1.7	3.3
6	12	9	6	3	2.5	0.8
7	13	9	10	7	1.5	3.9
8	20	24	13	4	5.0	2.5
9	27	27	23	15	1.1	2.2
10	24	19	15	10	1.5	5.0
11	12	9	10	5	2.0	2.2
12	28	27	18	8	5.0	11.0
Mean	21.3	21.1	14.6	9.2	2.8	4.0

Second Cycle

1	27	21	22	15	1.5	2.4
2	15	4	13	3	6.3	7.5
3	30	30	21	16	2.6	2.5
4	30	30	21	13	1.8	3.5
5	30	30	23	17	2.0	3.3
6	12	6	8	6	2.5	2.5
7	9	3	7	2	1.5	1.7
8	18	24	4	7	1.7	2.5
9	30	21	25	2	2.1	1.4
10	27	30	20	27	3.7	1.3
11	12	8	11	6	2.5	2.5
12	24	27	18	6	5.0	3.0
Mean	22.0	19.5	16.1	10.0	2.8	2.8
No. times superior	11	5	20	3	15	7
No. ties	8		1		2	

Total no. times *6 + 6* is superior = *46*
Total no. times *12* is superior = *15*
Total no. ties = *11*

Table VIII

Results of Experiment II.—Recall after 5 Minutes

First Cycle

Subjects	Total Digits		Correct Digits		Time per Digit (Seconds)	
	3 + 3	*6*	*3 + 3*	*6*	*3 + 3*	*6*
1	21	23	18	20	1.4	2.0
2	21	27	15	23	2.6	2.0
3	15	18	3	14	1.4	1.7
4	30	24	29	16	1.3	3.7
5	6	10	4	10	7.5	2.5
6	12	12	6	9	3.4	2.7
Mean	17.5	19.0	12.5	15.3	2.9	2.4
Second Cycle						
1	24	29	18	26	2.2	2.1
2	27	30	18	30	2.0	0.9
3	21	15	14	9	1.5	2.4
4	28	30	23	30	1.8	1.0
5	6	9	3	6	5.0	2.8
6	12	15	12	12	1.7	2.9
Mean	19.7	21.3	14.7	18.8	2.4	2.0
Third Cycle						
1	20	30	17	28	2.4	1.6
2	25	30	18	30	1.7	1.0
3	18	15	7	9	1.2	1.7
4	30	30	28	25	2.6	1.3
5	9	12	9	4	2.8	2.2
6	12	15	11	13	2.2	2.7
Mean	19.0	22.0	15.0	18.2	2.2	1.8

Fourth Cycle

1	30	25	25	22	1.6	2.0
2	27	30	24	20	1.6	2.2
3	21	12	14	7	0.8	0.6
4	30	30	27	30	2.3	1.1
5	6	6	3	5	7.7	6.1
6	12	12	3	7	3.2	2.9
Mean	21.0	19.2	16.0	15.2	2.9	2.5
No. times superior	5	14	7	16	9	15
No. ties	5		1		0	

Total no. times *3 + 3* is superior = *21*
Total no. times *6* is superior = *45*
Total no. ties = *6*

Table IX

Results of Experiment II.—Recall after 20 Minutes

First Cycle

Subjects	Total Digits		Correct Digits		Time per Digit (Seconds)	
	3 + 3	*6*	*3 + 3*	*6*	*3 + 3*	*6*
1	24	19	20	9	1.4	2.2
2	27	28	18	26	3.1	2.0
3	21	18	7	15	1.3	1.0
4	30	30	25	19	1.6	2.6
5	9	9	9	3	3.0	2.8
6	9	9	1	5	2.6	2.7
Mean	20.0	18.8	13.3	12.8	2.2	2.2

Second Cycle

1	20	25	17	21	1.9	3.4
2	25	30	22	23	1.4	1.5
3	18	24	6	22	1.5	1.3
4	30	30	30	28	1.3	2.0
5	9	9	4	4	1.9	2.7
6	9	15	4	7	2.1	2.7
Mean	18.5	22.2	13.8	17.5	1.7	2.3

Third Cycle

1	24	19	18	6	2.0	2.5
2	27	27	13	25	1.7	2.3
3	21	21	8	10	1.5	1.9
4	27	30	27	26	2.4	2.8
5	9	9	6	0	2.7	7.3
6	9	13	4	9	2.9	5.0
Mean	19.5	19.8	12.7	12.7	2.2	3.6

Fourth Cycle

1	21	24	19	15	3.0	2.2
2	24	27	21	27	2.0	1.4
3	21	18	13	15	1.7	0.9
4	27	27	24	27	3.0	3.1
5	9	6	6	4	3.9	5.2
6	9	6	9	4	2.5	6.3
Mean	18.5	18.0	15.3	15.3	2.7	3.2
No. times superior	6	9	10	13	17	7
No. ties	9		1		0	

Total no. times *3 + 3* is superior = *33*
Total no. times *6* is superior = *29*
Total no. ties = *10*

Table X

Results of Experiment II.—Recall after 24 Hours

First Cycle

Subjects	Total Digits		Correct Digits		Time per Digit (Seconds)	
	3 + 3	*6*	*3 + 3*	*6*	*3 + 3*	*6*
1	12	9	5	8	4.4	4.4
2	21	21	14	15	2.6	1.5
3	9	21	6	8	2.2	1.7
4	30	21	30	6	3.2	8.1
5	6	0	3	0	3.3	—[1]
6	9	9	7	1	2.2	1.8
Mean	14.5	13.5	10.8	6.3	3.0	3.5

Second Cycle

1	13	9	6	0	3.5	7.3
2	15	16	11	1	2.8	4.1
3	18	12	8	12	1.2	0.8
4	27	27	25	25	4.1	3.1
5	3	6	1	0	6.0	6.5
6	15	9	4	1	3.1	2.9
Mean	15.2	13.2	9.2	6.5	3.4	4.1

Third Cycle

1	17	25	8	5	3.2	3.5
2	18	18	9	0	2.7	3.7
3	18	21	9	8	1.4	1.2
4	24	27	7	23	5.1	3.1
5	3	0	2	0	21.0	—[1]
6	12	6	3	1	2.9	4.0
Mean	15.3	16.2	6.3	6.2	6.0	3.1

Fourth Cycle

1	11	12	5	8	3.2	3.7
2	15	12	6	8	3.9	4.8
3	21	18	14	12	1.5	1.6
4	22	24	20	15	4.8	5.8
5	0	3	0	2	—[1]	12.3
6	9	9	7	5	3.0	2.3
Mean	13.0	13.0	8.7	8.3	3.3	5.1
No. times superior	10	9	15	8	11	9
No. ties	5		1		1	

Total no. times *3* + *3* is superior = *36*
Total no. times *6* is superior = *26*
Total no. ties = *7*

From a similar pooling of all the comparisons of Experiment II., Tables VIII., IX., and X., it appears that:

Distributed study is superior 90 times
Concentrated study is superior 100 times
Neither is superior 23 times

These results strikingly support the generalization that the division of 12 presentations into halves 24 hours apart is highly advantageous, while a similar division of 6 presentations probably has a deleterious effect. Of course, a possibility remains that at least some part of this difference is due to the fact that different subjects were used under the two sets of conditions. Aside from this consideration, however, there are a number of factors which may operate to make distributed study relatively more advantageous in Experiment I. than in Experiment II.:

1. It is quite possible that the optimal time for putting an end to the first study of ten three-place numbers is after something more than 3 presentations;

2. Perhaps 3 presentations are fewer than are contained in optimal study periods for this material; and it is possible that:

[1] Subject no. 5 has no time values at these points because of inability to recall any digits.

3. If an interval of something less than 24 hours had elapsed between the first and last 3 presentations, the division of 6 presentations into halves would have been as advantageous as a similar division of 12 presentations.

Comparing the results of the two methods of study at different stages in the process of forgetting, it is evident that in Experiment I., Tables V., VI., and VII., the superiority of the distributed method of study is greater after 24 hours than it is after 5 minutes or 20 minutes. In the latter cases it is favored in 37 and 31 comparisons out of 72, and in the former case in 46 comparisons. This fact appears more plainly in these detailed data than in the averages of Table III.

The examination of Tables VIII., IX., and X., Experiment II., reveals facts in essential agreement with those of Experiment I. After 5 minutes, distributed study is favored in only 21 out of 72 comparisons, after 20 minutes it is favored by 33 comparisons and after 24 hours by 36. According to both of these experiments, then, the relative efficiencies of the two methods of study are functions of the time elapsing since the final presentations.

Again pooling the data of Tables V., VI., and VII., certain facts appear regarding the effects of utilizing different scoring methods:

According to the total number of digits recalled,

Distributed study is superior	*24 times*
Concentrated study is superior	*19 times*
Neither is superior	*29 times*

According to the number of digits recalled correctly,

Distributed study is superior	*44 times*
Concentrated study is superior	*21 times*
Neither is superior	*7 times*

According to the average number of seconds required to recall each digit,

Distributed study is superior	*46 times*
Concentrated study is superior	*23 times*
Neither is superior	*3 times*

Similarly, pooling the data of Tables VIII., IX., and X., it appears that, in Experiment II.:
According to the total number of digits recalled,

Distributed study is superior	*21 times*
Concentrated study is superior	*32 times*
Neither is superior	*19 times*

According to the number of digits recalled correctly,

Distributed study is superior	*32 times*
Concentrated study is superior	*37 times*
Neither is superior	*3 times*

According to the average number of seconds required to recall each digit,

Distributed study is superior	*37 times*
Concentrated study is superior	*31 times*
Neither is superior	*1 time*

Both experiments agree in indicating that *correct digits recalled* and *time of recall* show a greater advantage for distributed study than does *amount of recall.*[1]

The results of this investigation, then, suggest that the relative merits of distributed and concentrated study of numerical material depend upon: (1) the total amount of study considered, (2) the units into which that material is divided, (3) the stage in the forgetting process at which memorial efficiency is tested, and (4) the criterion of efficiency employed; *e.g.*, amount, accuracy, or time of recall.

Bibliography

Bergstrom, J. A. Effects of Changes in the Time Variables in Memorizing, together with Some Discussion of the Technique of Memory Experimentation. *Amer. J. of Psychol.*, 1907, 18, 206–238.

Browning, M., Brown, D. E., and Washburn, M. F. The Effect of the Interval between Repetitions on the Speed of Learning a Series of Movements. *Amer. J. of Psychol.*, 1913, 24, 580–583.

[1] There are many more perfect scores according to *amount of recall* than according to either of the other criteria—of course, a perfect time score would be as difficult to obtain as to define. Because of this fact there are more ties according to the amount criterion, and this naturally decreases its efficiency as a differentiating agency.

Carr, Harvey. Distribution of Effort. *Psychol. Bull.*, 1919, 16, 26–28.

Culler, E. A. The Effect of Distribution of Practice upon Learning. *J. of Phil., Psychol., Etc.*, 1912, 9, 381–383.

Cummins, R. A. Improvement and the Distribution of Practice. Teachers College, Columbia Univ. Contrib. to Educ., 1919, No. 97.

Dearborn, W. F. Experiments in Learning. *J. of Educ. Psychol.*, 1910, 1, 373–388.

Ebbinghaus, H. Ueber das Gedächtniss. Leipzig, 1885.

Hahn, H. H., and Thorndike, E. L. Some Results of Practice in Addition under School Conditions. *J. of Educ. Psychol.*, 1914, 5, 65–84.

Hunter, W. S. The Delayed Reaction in Animals and Children. Behav. Monog., 1913–1915, 2, No. 6, 30–31.

Jost, A. Die Assoziationsfestigkeit in ihrer Abhängigkeit von der Verteilung der Wiederholungen. Zschr. f. Psychol., 1897, 14, 436–472.

Kirby, T. J. The Results of Practice under School Conditions. Teachers College, Columbia Univ. Contrib. to Educ., 1913, No. 58.

Lashley, K. S. The Acquisition of Skill in Archery. Papers from the department of Marine Biology of the Carnegie Institution of Washington, 1915, VII, 105–128.

Lashley, K. S. A Causal Factor in the Relation of the Distribution of Practice to the Rate of Learning. *J. of Animal Behav.*, 1917, 7, 139–142.

Lashley, K. S. A Simple Maze: with Data on the Relation of the Distribution of Practice to the Rate of Learning. *Psychobiology*, 1918, 1, 353–367.

Leuba, J. H., and Hyde, W. An Experiment in Learning to Make Hand Movements. *Psychol. Rev.*, 1905, 12, 351–369.

Lyon, D. O. The Relation of Length of Material to Time Taken for Learning, and the Optimum Distribution of Time. *J. of Educ. Psychol.*, 1914, 5, 1–9, 85–91, 155–163.

Lyon, D. O. Memory and the Learning Process. Baltimore, Warwick and York, 1917.

Mould, M., Treadwell, L., and Washburn, M. F. The Influence of Suppressing Articulation on the Favorable Effect of Distributing Repetitions. *Amer. J. of Psychol.*, 1915, 26, 286–288.

Munn, A. F. The Curve of Learning in Studies in Development and Learning. Edited by E. A. Kirkpatrick. *Archives of Psychol.*, 1909, No. 12.

Murphy, H. H. Distribution of Practice Periods in Learning. *J. of Educ. Psychol.*, 1916, 7, 150–162.

Pechstein, L. A. Whole versus Part Methods in Motor Learning. *Psychol. Monog.*, 1917, 23, No. 99.

Pechstein, L. A. Massed *vs.* Distributed Effort in Learning. *J. of Educ. Psychol.*, 1921, 12, 92–97.

Perkins, N. L. Value of Distributed Repetitions in Rote Learning. *Brit. J. of Psychol.*, 1914, 7, 253–261.

Pyle, W. H. Economical Learning. *Psychol. Bull.*, 1913, 10, 73.

Pyle, W. H. Economical Learning. *J. of Educ. Psychol.*, 1913, 4, 148–158.

Pyle, W. H. Concentrated *vs.* Distributed Practice. *J. of Educ. Psychol.*, 1914, 5, 247–258.

Starch, D. Periods of Work in Learning. *J. of Educ. Psychol.*, 1912, 3, 209–213.

Strong, E. K. Two Factors which Influence Economical Learning. *J. of Phil., Psychol., Etc.*, 1914, 11, 124–131.

Strong, E. K. The Factors Affecting a Permanent Impression Developed through Repetition. *J. of Exp. Psychol.*, 1916, 1, 319–338.

Thorndike, E. L. Notes on Practice, Improvability, and the Curve of Work. *Amer. J. of Psychol.*, 1916, 27, 550–565.

Ulrich, J. L. Distribution of Effort in Learning in the White Rat. Behav. Monog., 1915, 2, No. 5.

Yerkes, R. M. The Dancing Mouse. New York, Macmillan, 1907, 239–250.

THE COLORS PRODUCED BY EQUILIBRIUM PHOTOPIC ADAPTATION

BY LEONARD THOMPSON TROLAND,

Harvard University, Emerson Hall

I. Introduction

In a recent paper Sheppard reports extensive and seemingly careful experiments which tend to support Hering's belief—a necessary deduction from his visual theory—that equilibrium adaptation of the retina to any stimulus yields a sensory quality of neutral mid-gray. Sheppard[1] finds that with practically all types of stimuli, whether high or low in intensity and whether or not of spectral purity, the final result of continued fixation is a gray. Sheppard's results seem to be in definite disagreement with the general outcome of quite an elaborate series of experiments which the writer made in 1914-15 at the Harvard Psychological Laboratory and which have not hitherto been described in detail. The experiments led to the conclusion that prolonged stimulation of the retinal cones with stimuli which initially evoke colors of considerable saturation does not in general bring about reduction of the sensation to a gray. In certain instances, the conditions of which can be quite definitely specified, such a reduction does occur, but the general rule was in sharp conflict with the implications of the Hering theory.

The experiments were reported verbally in summary at the meeting of the Experimental Psychologists at Harvard in 1917, and it is possible that Sheppard's work may have some indirect reference to the views which were advocated at that time. The results have also been mentioned in a general paper on "Apparent Brightness,"[2] incidentally in a paper on

[1] Sheppard, H., 'Foveal Adaptation to Color,' *Am. Journ. of Psychol.*, 1920, 31, 34-58.

[2] Troland, L. T., 'Apparent Brightness: Its Conditions and Properties,' *Trans. of the Illum. Eng. Soc.*, 1916, 11, 949-966.

special after-image phenomena,[1] and were utilized in 1914 to substantiate a general mathematical theory of adaptation[2] which was based on assumptions contrary to those of Hering. Since the question which is at issue is of great importance in determining the value of Hering's general conceptions for visual theory, as well as in relation to any supposed mechanism of visual excitation, it seems desirable that the results be reported in full. The writer believes that the discrepancy between his observations and those made by Sheppard and other supporters of the Hering theory will be found in some difference between the conditions under which the experiments were made, the principal factor being in all probability the control of the pupil. The majority of Sheppard's observations were made with the natural pupil, and those in which he employed an artificial pupil apparently did not involve any accurate control of the registration of the artificial pupil with the natural pupil. In all of the observations reported in this paper, on the other hand, or at least in those upon which reliance is placed for the refutation of the Hering theory, an artificial pupil was employed and care was taken to secure satisfactory register between the natural and artificial pupils. It is worthy of note that in Sheppard's series with spectral stimuli[3] in which an artificial pupil was employed he found it impossible to obtain 'complete adaptation' at high intensities.

The spontaneous fluctuations to which the natural pupil is almost always subject, produce variations in the illumination of the retina, and such variations result in very powerful after-image effects which are capable of quite submerging the normal course of the adaptation process. Only by carefully eliminating such fluctuations, and their accompanying changes in retinal illumination, can we hope to study the adaptation process in its purity. One of the most interesting features of the writer's work on foveal adaptation, only a portion of which will be reported in the present article, lay

[1] Troland, L. T., 'Preliminary Note: The Influence of Changes of Illumination upon After-Images,' *Amer. J. of Psychol.*, 1917, 28, 497–503.

[2] Troland, L. T., 'Adaptation and the Chemical Theory of Sensory Response,' *Amer. J. of Psychol.*, 1914, 25, 500–527.

[3] *Loc. cit.*, p. 55.

in the quantitative study of these secondary after-image phenomena which appear with changes in the retinal illumination. The phenomena in question may consist either in a marked augmentation of the effect which goes with stationary illumination or, on the other hand, they may constitute a reversal of the later effect. The first type of change goes, in general, with diminution of the retinal illumination, caused by contraction of the pupil or otherwise, while the second species of influence accompanies the reverse change. It is true that in general these illumination-change phenomena are short lived; but they are of sufficient duration to cause an observer to make a judgment of the consummation of complete adaptation and thus bring a given observation to a close.

Some of the general principles which determine one of these effects have already been discussed briefly in another paper, and further details will be considered in the present and later articles.

II. The Quality of the Equilibrium Sensation

The majority of theories of visual response assume that the intensity of any given component of the visual sensation is a function of the velocity of a corresponding chemical change. With the possible exception of the sensation of black, this function is such that when the velocity in question is zero the intensity of the corresponding sensory *quale*[1] is also zero, and also such that an increase or decrease in the velocity is characteristically accompanied by an increase or decrease, respectively, in the intensity. Since it is customary to represent chemical velocities as rates of change in the concentration of the reacting substance, or substances, we may state that if the visual intensity considered is v, while the concentration of the substance is s, and t the time, these theories require in general that:

(1) $$v = f^d(\pm\, ds/dt),$$

where $f^d(\ )$ is a function having the properties specified above.

The above general premise of all chemical theories of vision has obvious, important consequences, on account of

[1] By "the intensity of a sensory quale" is meant the measure of any attribute, *e.g.*, the degree of redness or of brightness.

the inevitable law that the velocity of a chemical change decreases or increases with a corresponding alteration in the concentration of the active substances, and such a change in concentration is the very thing called for by the rate of change itself. In general, then, without undue exactness of expression:

(2) $$ds/dt = -f^d(s)$$

and

(3) $$s = \int(ds/dt)dt,$$

that is, other things equal, the velocity is greater the greater the concentration of the reacting substance,[1] and the concentration is greater the longer the time during which the change has been occurring and the higher the velocity during that time. If we limit ourselves to a so-called irreversible monomolecular reaction, equation (2) becomes a proportionality, so that

(4) $$ds/dt = -ks,$$

where k is a constant. This equation states that the rate at which the substance, S, is being destroyed is at every instant proportional to its concentration. A simple integration of (4) gives us the relationship:

(5) $$\log s = -kt, \quad \text{or} \quad s = e^{-kt},$$

where e is the base of the natural system of logarithms. From (4),

(6) $$ds/dt = -ke^{-kt},$$

which asserts that the rate of decomposition of S becomes less as time goes on, approaching zero 'asymptotically.' But, from (1) it follows that

(7) $$v = f^d(\mp ke^{-kt}),$$

which states that with time the sensation *quale*, V, becomes less and less intense, and after a sufficiently long period will be reduced below the threshold and hence, psychologically, to

[1] The minus sign in equation (2) is due to the convention of taking the rate of change as negative when it is destructive of the substance, S.

zero. This process obviously corresponds to what is usually called visual or retinal 'fatigue,' and in a general way the actually observed changes in v with time follow a curve similar to that demanded by the above equation (7).

However it is one of the principal points of the experimental work which is to be described below to show that, at least in one respect, equation (7) is not generally valid, *viz.*, that it is not an essential characteristic of visual response that continued activity should result in ultimate complete exhaustion. Now all chemical theories of vision provide a mechanism of 'repair' whereby the change in the concentration of the sensitive substance which occurs during sensation can be compensated. If the original change was destructive or 'katabolic,' the repair is of course constructive, or 'anabolic,' and the latter change always implies a velocity term opposite in sign to that of the former. It is generally assumed that the repair goes on during stimulation as well as thereafter, and if this is the case it is obvious that the concentration of the sensitive substance will never be reduced to zero, no matter how long the excitation continues. If we suppose the repair process to have a velocity r at any time the intensity of the visual *quale* at that time will be

$$v = f^d(\mp ke^{-kt} \pm r), \tag{8}$$

so that after a very long time:

$$v = f^d(\pm r). \tag{9}$$

If r is a constant or becomes so after the 'long' period of time, equation (9) represents a steady state of sensation which will continue indefinitely. Such a condition, in general, we shall call one of *sensory equilibrium.*

It is characteristic of the visual theory of Hering[1] that certain visual qualities are attributed not only to the katabolic but also to the anabolic phase of the metabolism, so that the stimulation process of one sensation is the repair process of another sensation, which is defined as its 'antagonist.'

[1] Hering, E., 'Zur Lehre vom Lichtsinne,' *Sitzb. d. Wien. Akad.*, 1872–1874. 2 *Aufl.*, 1878, p. 70.

Moreover, except in the case of the 'black-white substance,' when the opposed velocities are arithmetically equal these paired qualities completely inhibit each other, so that the theory in question implies a complete absence of both of these antagonistic qualities under conditions of sensory equilibrium.[1] As is well known, the qualities thus opposed are red to green, and yellow to blue. It is a recognized inconsistency of Hering's theory that black and white, although given the same psycho-physical basis as the other visual qualities, do not antagonize, but instead fuse with each other.

In harmony with Hering's point of view and what is in all probability a correct psychological analysis, we shall consider any visual sensation as resolvable into a limited number of component intensities, *viz.*, different *degrees* of black, white, red, yellow, green and blue. For these qualities we shall employ the abbreviations: N (noir), W, R, Y, G, and B, respectively. For the corresponding degrees we may employ the lower case letters: n, w, r, y, g, and b. Everyday experience teaches that, ordinarily at least, R and G are mutually exclusive, as are also Y and B; but that with this restriction a visual sensation may exhibit any combination of the variables named, with a wide range of relative intensities.

Let us define as an *equilibrium sensation* any sensation which accompanies a process in the visual system which involves no concomitant alteration in the sensitivity of the latter. Such a sensation will remain when the process of fatigue or adaptation has reached its asymptotic limit.

Now, as we have stated, it is a consequence of Hering's theory that there is only one possible constitution for such a sensation, *viz.*, $N+W$, in which $w=n$, a mixture denominated 'mid-gray.' According to this view, it can make no difference whether the equilibrium sensation is determined by an external or an internal stimulus, or whether the original sensation was chromatic or achromatic. All sensory processes, if permitted to go on undisturbed, must finally meet a common fate.

The most obvious method for an attempt to refute this consequence of Hering's theory would seem to be to fixate

[1] Hering, E., 'Grundzüge der Lehre vom Lichtsinn,' 1905–1911.

different stimulus fields during long periods, and to observe whether a time can be reached when all of the differentiæ resolve themselves into a mid-gray. In the present paper a report is made of the results of experiments of this character, the stimuli employed covering a fairly wide range of conditions. The abolition of all visual differences during a period of fixation amounts to a 'disappearance' of the stimulus, and it is a fact so well established that it must be accepted at the outset that such disappearances often occur. Hence one of the problems to be considered is that of the exact circumstances which make such lapses possible, and whether they argue for Hering's view. This is a field of investigation in which already a great deal of work has been done, but in which, nevertheless, the writer believes he has found certain new facts. The final outcome of the paper will be a denial that the equilibrium sensation is a neutral mid-gray, except in certain very special cases, and the phenomena of stimulus disappearance will be given explanations which do not imply this proposition.

III. The Equilibrium Sensation with Large Uniform Stimulus Fields

The simplest case in which to attack the problem of the quality of equilibrium sensations would seem to be that of the complete exclusion of light from the normally conditioned eye. This cannot be called an absence of stimulation, since, especially in such a view as Hering's, it is proper to consider darkness as a definite visual stimulus. But if Hering's doctrines are correct, darkness, like all other environmental conditions, is a stimulus only when it is a change. Black results from the cutting off of light, but the continued absence of light is accompanied by a lapse of the black sensation into a neutral mid-gray. The question therefore presents itself for empirical consideration as to whether the sensation yielded by the completely rested retina in the dark is actually neutral mid-gray. Hering himself considers this point at some length[1] and admits that the visual quality which one experi-

[1] 'Zur Lehre vom Lichtsinne,' p. 89.

ences upon awakening in the morning in a closely shuttered room suggests black more than it does white, and hence might be thought not to be such a gray. Hering explains this appearance by saying that although we have had experience of practically perfect whites, we have never had one of absolute black and that, consequently, our conception of mid-gray is weighted in the direction of white. This makes a true mid-gray seem too dark.

The author's observations confirm the naïve impression that the rest idio-retinal sensation is a great deal blacker than an unsophisticated mid-gray, but his reasoning does not wholly substantiate Hering's attempted reconciliation of such observations with his theory. In the first place, it is doubtful whether we can come any closer to an experience of pure white than we can to one of pure black. What reason have we for believing that 'pitch dark' lies farther from utter blackness than 'dazzling white' does from a catholic whiteness? Hering's argument, to say the least, presupposes his own theory. If, as such hypotheses as those of Helmholtz would imply, black is a result of the absence of retinal excitation, it would seem easier to produce a pure black than a pure white, since, theoretically, the white process could increase indefinitely in intensity whereas the black process would find its natural limit with complete retinal fatigue, a conceivably attainable state of affairs. In accordance with Hering's hypothesis, on the other hand, no natural limit can exist for either n or w. Consequently it would appear necessary to estimate both blackness and whiteness on the basis of mid-gray as a standard, rather than *vice versa*. If this is what we actually do the empirical evidence would seem to be against Hering's doctrine of the median quality of the equilibrium sensation in the absence of external stimulus forces.

As will be emphasized more strongly later on, it is probable that there is a central process in vision—closely allied with the functions of perception and discrimination—which causes us to become 'mentally blind' to sensory qualities which are uniformly distributed in both space and time. This process is slow-acting, but it obeys a law similar to that governing Hering's equilibrium sensation: it makes it impossible for an observer to 'tell' whether or not the monotonous quality is present, and hence favors judgments of neutrality with regard to all sensations which have persisted for a long time unchanged. If such a process is active in the limitation of perception and introspection, we should expect the idio-retinal rest sensation to be judged as mid-gray rather than as black or dark gray. Thus, if central adaptation fails to take effect, surely there must be some adequate peripheral cause, such as marked and permanent depression of the retinal activity below that characteristic of mid-gray.

Suppose, now that in place of darkness we fill the eye with a bright white light, and continue such stimulation indefinitely. If Hering's theory is correct, no matter how brilliant the light the final sensation will be that of neutral mid-gray.

Experiment I. The Equilibrium Sensation for Direct Sunlight on White Paper

As a preliminary test of this point the following experiments were tried. A large sheet of white drawing paper directly illuminated by the early afternoon sun on a cloudless day was viewed at such a distance as to yield a uniform stimulation field of 60 horizontal and 50 vertical degrees. Fixation of this field, which had a brightness — calculated from direct flicker photometric measurements—of about 346,600 candles per square meter, was steadily continued for ten minutes, in each trial, with the following results.

At the outset the stimulus appeared dazzlingly bright, and during the first thirty seconds or so it showed a very rapid decrease in apparent luminosity, so that at the end of about one minute it seemed as if illuminated by, roughly, one twentieth of the actual light intensity. Whatever significance may be assigned to an estimate of this sort, it is certain that the gray sensation to which it refers was very much brighter in each case than a naïvely conceived mid-gray. Moreover after the first minute no further darkening could be observed, and during the ensuing nine minutes whatever alteration did occur was in the direction of a restoration of the original brightness rather than that of its continued decrease. This experiment was repeated a number of times on different days with the same results. It would appear from the above observations that (1) prolonged exposure of the retina to large, very intense, achromatic stimuli does not involve a reduction of the sensation to neutral mid-gray, but rather to a *bluish* gray the luminosity of which depends upon that of the stimulus and is much greater than that of mid-gray; and (2) that the reduction in luminosity which does take place occurs very rapidly so that the sensation soon comes to a permanent equilibrium state characterized by the just mentioned quality.

Experiment II. The Equilibrium Sensation for Direct Sunlight on Large Colored Papers

The principle of the reduction of visual sensation, by adaptation, to neutral mid-gray applies in Hering's theory not only to black and white but to all of the color qualities as well. As already stated, this is a consequence of his doctrine of mutually antagonistic color pairs. To put the question to an empirical test, large sheets of colored paper (of the Hering and Milton Bradley series) were fixated in direct sunlight under conditions exactly similar to those obtaining for the white stimulus just considered. Table I. summarizes the results of these experiments.

It will be seen by reference to the table that only in the case of yellow and green was there total disappearance of the color quality which was characteristic of the primary field and, even here, the disappearance was not permanent. In all cases there was, of course, a marked decrease in saturation and brightness, and sometimes a change in the exact hue-constitution of the field, very soon after the exposure was commenced. This change did not continue, however, and the quality of the area quickly settled down to an equilibrium condition which—apart from some tendency to revert to the original hue—was steadily maintained thereafter. It should be noted that those cases in which the primary hue showed phases of complete disappearance, were with stimuli of very high intensity, both the Milton Bradley yellow and Hering green being papers of high reflecting power.

Table I

The Fatigue Changes in Sensation with Large Chromatic Stimuli of High Intensity

General Conditions: Size of stimulus fields: approximately 60° horizontally by 50° vertically; Time of each fixation period: 10 minutes. Colored papers from the Hering (Rothe) and Milton Bradley series were employed. These were viewed in direct sunlight on cloudless March and April days, between 1 and 3 P.M. Subject: *T.* Monocular fixation.

Explanation of Tables: In the first column, below, are symbols for the normal quality of the stimulus, together with the abbreviations, Her. and M-B. to indicate whether the papers used belong to the Hering or to the Milton Bradley series, respectively. The second column given the calculated (approximate) value in candles per square meter of the brightness of each of the paper surfaces, the data for these calculations being obtained by flicker photometry of the papers actually employed. The next three columns contain descriptions of the quality of the sensation during three successive periods of the fixation, the first being that introduced by the rapid initial change, the third the final minute, and the second an intermediate phase.

Stimulus	Intensity	Color Phases			Remarks
		Initial	Intermed.	Final	
NR M-B.....	38,500	*r*	*r*	*r*	
R Her.....	50,000	*wr′*	*wr*	*w′r*	Clear revival of the red
YR Her.....	88,500	*wbg*	*yg*	*wry*	
Y M-B.....	215,000	*w*	*wgy*	*wgy*	Yellow very faint
Y Her.....	130,000	*gy*	*g′y*	*g′y*	
G Her.....	80,500	*wg′b′*	*wgb*	*wgb*	Moments of gray
GY M-B.....	117,500	*b′w*	*b″g″w*	*b″g″w*	Green very faint
BG Her.....	31,000	*b′w*	*b′w*	*b′w*	
B M-B.....	69,000	*w″b*	*w″b*	*w″b*	Little change
RB Her.....	37,500	*w″rb*	*w″rb*	*w″rb*	Little change

Method of Symbolism: The letters used in the above table refer roughly to the component intensities defined at the beginning of the present paper. The 'primes' are intended to indicate a *weakness* of the primed component *relative* to the others. Thus *gy* stands for a quality about equally green and yellow, *g′y* for one in which the yellow is considerably stronger than the green, and *g″y* for one in which it is far stronger. The quality orange is naturally symbolized by *ry*, and purple by *rb*. The component *n* (degree of darkness) is neglected in all cases, and *w* is used only when needed to indicate a lack of saturation greater than that normal to the given stimuli. Strictly speaking, of course, both *N* and *W* are components of all sensations, whether chromatic or achromatic.

General Results: The table shows that in those cases where the primary quality of the stimulus entirely disappears or becomes very weak, the intensity is greater than 50,000, and frequently than 100,000 candles per square meter. The stimuli which are weaker than 50,000 candles per square meter show relatively little change. The already familiar fact that *B* is the most stable of the color qualities is again revealed in this table.

The following generalizations may safely be made from the above described experiments. (1) Prolonged exposure of the retina to large chromatic stimuli of high, but not maximal, intensity is not accompanied by the disappearance of the hues characteristic of those stimuli or, necessarily, by a reduction of the brightness value of the sensation to that of a neutral mid-gray. (2) With such stimuli the qualitative change which does occur occupies a comparatively short time at the beginning of an exposure, after the end of which time a qualitative equilibrium becomes established. (3) When the chromatic stimuli employed are of exceedingly high intensity, complete disappearance of the characteristic hues may ensue. These disappearances occur soon after the beginning of an exposure, but they do not necessarily involve the abolition of all chromatic quality, and never mean a reduction of the sensation to the achromatic level of neutral mid-gray.

The preliminary experiments just described suggest very strongly that the equilibrium sensation in vision is not a monotonous gray but, instead, varies widely with the nature of the stimulus in connection with which the equilibrium is established. Of course we cannot be certain at the outset just how long it should take to bring the visual system to equilibrium under given conditions, but there are reasons for believing that in general the time should not be very great. The above experiments show that the sensory changes which follow the onset of stimulation at first involve a quite rapid decline of the characteristic stimulus quality; but although the alteration in the sensation during the first 30 or 60 seconds of the fixation is very readily perceived, that which takes place during the second minute (say) is—with ordinary light intensities—so small as to be hardly perceptible. Now it is just this sort of change which is demanded by a curve of the type presented in our equations (7) and (8) which, as we have seen, must go with such a chemical theory of vision as Hering's. However, there is a further qualitative consequence of the applicability of this form of equation to the phenomena which is not in harmony with Hering's teachings. If the changes in sensation during a given constant interval of time are at first large, but rapidly become small, the total change occurring thereafter prior to the attainment of practically complete equilibrium, will also be very small compared with the initial changes. In other words, if for the consummation of the greater part of the change in sensitivity of the organ towards equilibrium a long time is to be assumed, it must be supposed also that a long time will be required for any alteration in the *rate of change* of sensitivity to occur. The mathematical explanation of this is apparent when it is noticed that all of the derivatives of the function $-ke^{-kt}$ (equation 7) with respect to t are equal to the function itself, and hence to each other.

If follows from this that, neglecting for the moment the effect of the function $f^d(\ )$ upon the changes, the fact of a rapid initial decrease in v, followed by a very slow decrease, indicates that the sensation reached in this brief preliminary period of fatigue is very close to the equilibrium sensation. Observation shows that in the majority of cases the approxmate equilibrium sensation thus obtained is very far from being a neutral mid-gray. Moreover, greatly prolonged fixation, as we should expect, brings about relatively little alteration in the quality which is left after the first few minutes of fatigue. As already noted, there are strong indications that, if anything, continued fixation may entail a reversal in the direction of the sensitivity change. In this event the fatigue is limited not by an asymptote but by a maximum, and the actual law of the adaptation process must involve functions either different from or in addition to that of equation (8). We shall return to this problem at a later point in the discussion.

With regard to the interpretation of the above results as a refutation of Hering's assumptions regarding the condition obtaining at visual equilibrium, several objections may be raised, some theoretical, some experimental. They all tend to throw doubt upon the notion that actual conditions of equilibrium were reached in the experiments described.

In the first place it may be asserted that the form of the function, f^d ( ) in equations (7) and (8) is such as to magnify changes in $-k\,e^{-kt}$ for large values of the latter and to minify such changes for small values, so that the time for the attainment of approximate equilibrium may really be great, although the apparent changes in the sensation are larger during the earlier than during the later parts of the fixation period. In answer to this it may be said that there can hardly be any doubt that f^d ( ) is logarithmic in general form—as demanded by Fechner's law relationship between the physiological visual process and the sensation — and that, consequently, its effect will be just the reverse of that required to weaken our argument against Hering. Equal changes in $-k\,e^{-kt}$ will produce smaller variations in v for high than for low values of the function in question.

Another objection, based upon experiment, may be drawn from the fact that if colored glasses are worn for a relatively long period of time, or if one is shut up in a room illuminated by a single color, one eventually becomes quite blind to the color in question. The experiment of Maria Bokowa[1] is classical in this connection. However, as previously indicated, there are strong reasons for believing that, besides visual fatigue in the ordinary sense of the term, we have in such phenomena as these to deal with a type of 'higher,' central alteration in sensation, such as is involved in the study of so-called 'memory colors.' When we are placed under the practical necessity of discriminating between objects for purpose of action we perceptually neglect all absolutely common qualities of these objects. Such perceptual neglect, or central adaptation may result, with sufficiently long exposure to a single color, in a nervous set which will remain for some time after the color is removed. This would involve blindness to the quality in question. The existence of a positive central control over the visual sensations cannot be denied.

[1] Bokowa, M., 'Ein Verfahren künstliche Farbenblindheit hervorzubringen,' *Zsch. f. rat. Med.*, 1863, 17, 161 ff.

IV. The Influence of Possible Central Factors on the Perception of Chromatic Qualities at Equilibrium

That central processes are actually involved in the fading of constantly fixated uniform color fields of large area is strongly suggested by the following observations.

Experiment III. The Restoration of a Chromatic Quality, Faded Through Adaptation, by the Introduction of a Neutral Comparison Object

If, during the intermediate period of the fixation of any of the large colored papers in Experiment II., a small white or gray object—such as the dial of a stop-watch—is introduced into the field, the original color is perceived with suddenly increased distinctness. The alteration in the appearance of the field is very remarkable, and yet it does not seem to be a purely sensory change: it lacks the characteristic *quale* of an objective alteration in illumination conditions and feels more like the reëstablishment of a decadent perception. This effect occurs for all the colors employed in Experiment II., regardless of the stimulus intensity. When the comparison object is introduced the color field turns darker at the same time that its original quality is again clearly distinguished.

In these experiments the comparison object is of course suffused with a color complementary to that of the original field, and hence the rejuvenescence of the latter quality would naturally be attributed to simultaneous chromatic contrast. It is clear, however, that with regard to the retina at least, the situation must be somewhat different from the usual case in which the contrast-inducing color depends upon the specific quality of the stimulus. Moreover, in the effect which we are now considering the contrast-inducing object is small compared with the field which it affects, and yet its influence is very striking. Its law would therefore appear to be quantitatively, if not qualitatively, different from that of ordinary color contrast. That the two cases are really distinct is proven by the following experiment. A dark gray square of size approximately 9 degrees, was placed in the center of the yellowish green Hering paper, and was fixated in direct sun light until the green component had entirely faded, leaving a light blue—a state of affairs which continued for four minutes. At the end of this time a shutter arrangement was released which substituted in place of the gray square an equal area of Hering violet paper. Under these conditions there was no reappearance of the green, the outlying field becoming simply a somewhat darker blue. Immediately, however, a white object was introduced into the green adapted area, the latter color returned.

Simple tests were made to determine whether the reappearance of the adapted color might not be due to a contraction of the pupil, owing perhaps to the correction of a relaxed accommodation when the comparison object was introduced. In these tests an object of about the same brightness as the large field was employed, and care was taken to keep the fibers of the paper clearly in focus throughout. This made no difference in the results. As will appear later, it is probable that contraction of the pupil would tend to bring out the complementary quality rather than that proper to the field, although it is not entirely certain that this would be the case for the high degrees of illumination here employed.

The above experiments, while they support the view that color sensation is subject to central control, also indicate that 'central adaptation,' if it be a fact, can function as a local process, and hence is not strictly equivalent to 'forgetting what the color looks like.'

Experiment IV. The Equilibrium Sensation for Large Contrasting Colored Paper Fields

In order to study the importance of contrast and discrimination processes for the changes occuring in large fixated fields of color, observations under conditions similar to those of Experiment I., were made; the stimuli consisting of two juxtaposed Hering papers of different hue, the one filling the upper half of the field and the other the lower half. The six possible combinations of the four *Urfarben* were employed, and the results are shown in Table II.

An examination of the table in question will reveal the presence of a rather well marked mutual influence of the contrasting colors of the stimulus in the determination of the equilibrium qualities. This interaction appears to follow the law of ordinary complementary induction, as would be expected. For example, the red with the green loses less of saturation than it does with the yellow, while with the blue it becomes orange instead of pink. The green with the blue becomes yellowish, whereas with the yellow it turns toward the blue. The cases in which this complementary influence is smallest appear to be with the blue and with the yellow. This fact may be attributed to the much greater brightness of the yellow and greater darkness of the blue, as compared with the other two colors, or else to the well-known superior stability of the *Y* and *B* components, in general. The results of the present experiment would appear to support the conclusion that the principle of simultaneous color contrast applies to the equilibrium condition of the sensation, as well as to other phases. Whether this fact is to be attributed to central or to peripheral action is a question which must be considered an integral part of the still unanswered query as to the meaning of contrast, at large. From the writer's present point of view it seems to him probable that complementary color contrast in general is a central process, being one case of the sort of regulative activity of the centers with regard to the sensation quality already suggested above.

It must not be inferred from what has just been said that the results summarized in Table II. represent a contrast effect of the same order of magnitude as that considered in Experiment II. On the contrary, the introduction of a neutral object into the bi-chromatic field brings about the same striking change in quality, for both halves of the field, which occurs in the case of the simpler stimulus.

Observations under conditions exactly similar to those specified in Table II. were carried out with bipartite fixation fields composed of black and white—black velvet and white drawing paper—respectively. The intensity of the former was about 1,250 candles per square meter, and of the latter

Table II

The Fatigue Changes in Sensation with Large Bichromatic Stimulus Fields of High Intensity

General Conditions: Size of stimulus fields: approximately 50° by 50°, the upper half being filled by one color and the lower by the other; Time of each fixation of the middle of the line of demarcation: 10 minutes. The papers were from the Hering series only. These were viewed in direct sunlight on cloudless March and April days, between 1 and 3 P.M. Subject: *T*. Monocular fixation.

Explanation of Tables: The arrangement of the table, and the symbolism employed is the same as for Table I., above. Brackets are used to indicate the pairs of stimuli fixated together in the double fields.

Stimulus	Intensity	Color Phase: Initial	Intermed.	Final	Remarks
{ *Y*	130,000	*y″g″w*	*ygw*	*ygw*	There was a short initial phase of nearly equal luminosity.
{ *R*	50,000	*r″w*	*rw*	*rw*	
{ *R*	50,000	*r′w*	*rw′*	*rw′*	
{ *G*	80,500	*w*	*g″b″w*	*g″b″w*	
{ *R*	50,000	*ryw*	*ryw*	*ryw*	
{ *B*	6,700	*bw*	*bw*	*bw*	
{ *Y*	130,000	*gy*	*g′y*	*g′y*	Same effect as for *Y* and *R*, above.
{ *G*	80,500	*b′w*	*b′w*	*b′w*	
{ *B*	6,700	*bw*	*bw*	*bw*	
{ *Y*	130,000	*gy*	*gy*	*gy*	
{ *B*	6,700	*bw*	*bw*	*bw*	
{ *G*	80,500	*w*	*yw*	*yw*	

approximately 250,000. At no time during the ten-minute fixation period was there even a tendency for the two halves to fuse in a common neutral mid-gray, as demanded by Hering's theory. The white became greatly reduced in apparent luminosity and assumed a bluish tinge during the first minute, but after that suffered no change, except possibly a slight increase in brightness. The black showed no change at all. Experiments were also made with a piece of white paper bearing a diffuse smudge of smoke—a pattern recommending by Hering for the demonstration of the lapse of visual differentiæ through fatigue[1]—which was fixated in bright sunlight. Momentary disappearances of the brown smoke spot were obtained, but these were always during the first minute or two after fixation. Even at the outset the

[1] Hering, E., 'Zur Lehre vom Lichtsinne,' *Sitzb. d. Wien. Akad.*, 1872–1874, 2 Aufl., 1878, p. 96.

smudge was a fairly bright gray, a fact which helps to account for its easy disappearance.

In all of the experiments thus far described the writer acted as subject.

The experiments described in the present Section offer a certain degree of confirmation for the idea suggested at the end of Section I. that the disappearance or weakening of some of the components of visual sensation may have central as well as peripheral causes, a procedure being pointed out whereby the original qualities can be restored without resting the retina. The success of this procedure appears to rest upon a central principle of color contrast. In line with this it is shown that the law of complementary color induction applies to the equilibrium, as to other, visual sensations.

V. The Fluctuation in Visibility of Spot Stimuli on a Dark Ground, and Related Phenomena

Suppose, now, that we pass from the consideration of the probable equilibrium conditions which go with large intensive stimuli of uniform quality to those obtaining with simple patterns in which the illumination is less intense. The simplest case is that of a small round luminous spot on a dark ground. A large number of investigations are already on record with regard to the course of the sensation conditioned by such stimuli, with constant fixation. It is a well-known fact that when the illumination of the spots is at low intensity they readily disappear, and either remain below the threshold or fluctuate between visibility and invisibility.[1] If the dark outlying field could be considered as a neutral mid-gray, this would present a case of the complete reduction of a sensation by adaptation. These disappearances also occur when the intensity of the stimulus is moderately high, but they are less frequent and last a much shorter time than with the dim lights. The general law appears to be that the phase of visibility increases at the expense of that of invisibility with increasing intensity.[2] For very low intensities disappearance may be

[1] Ladd, G. T., 'A Color Illusion,' *Yale Studies*, 1898, 6, 1 ff.

[2] Marbe, K., 'Die Schwankungen der Gesichtsempfindungen,' *Phil. Stud.*, 1893, 8, 615–637.

permanent, if the spot is not too small, but for very high ones it never occurs in the first place.

These fluctuations of visual sensation under the influence of a constant external stimulus receive ready explanation in terms of the Hering theory, if it is supposed that slips of fixation occur, during which a rapid change in the state of adaptation of the visual system can take place. If, prior to the eye-movement, the visual cells were at equilibrium with the stimulus, the field of sensation must have been a neutral mid-gray, but as soon as the ocular displacement occurs, differential recovery sets in so that upon a return to the original fixation the spot is again momentarily seen. Ferree[1] has reported experiments in support of this interpretation of the fluctuations. The Hering theory of the mid-neutrality of the equilibrium sensation, however, does not account for the relation of the relative lengths of the phases of visibility and invisibility to the intensity of the stimulus, although it is directly in harmony with the fact, adduced by Ferree, that for certain low intensities very small areas show no disappearances while very large ones disappear permanently.[2] The imperfectness of fixation always suffices constantly to present fresh retinal surfaces to a very small stimulus spot. For complete harmony with the theory, the rule of disappearance, without return, of large areas should hold for all intensities.

The fluctuations in the gray bands presented by the classical Masson's disk have, for a long time, been attributed to momentary 'lapses of attention,'[3] in spite of the obvious fact that under the usual conditions of the experiment, attention to the disk is probably maximal at just the moments of disappearance of the bands. Heinrich and Chwistek,[4] and others, have sought the cause of the fluctuations in spon-

[1] Ferree, C. E., 'An Experimental Examination of the Phenomena Usually Attributed to Fluctuation of Attention,' *Amer. J. of Psychol.*, 1906, 17, 81–120.

[2] Ferree, C. E., 'The Intermittence of Minimal Visual Sensations,' *Amer. J. of Psychol.*, 1908, 19, 59–129.

[3] Lange, N., 'Beiträge zur Theorie der sinnlichen Aufmerksamkeit und der activen Apperception,' *Phil. Stud.*, 1888, 4, 390–422.

[4] Heinrich, W. und Chwistek, L., 'Ueber das periodische Verschwinden kleiner Punkte,' *Zsch. f. Sinnesphysiol.*, 1906, 41, 59–74.

taneous changes in lens accommodation, a quest which appears in the light of later investigations to have been in vain.[1]

Experiment V. The Dependence of the Time of the First Fluctuative Disappearance upon the Intensity of the Stimulus

That these fluctuative disappearances of sensations caused by constant stimuli actually have the adaptation process as at least a partial basis is strongly suggested by the following experiment. A white circle, of visual angle approximating 4°, was fixated monocularly in the dark (by subject *T*) and the time from the beginning of the fixation until the occurrence of the first fluctuation was measured. When the light was at low intensity the average of four trials gave 15 seconds (A.D. 1.6) for the first disappearance, and 24 seconds (A.D. 1.9) for the second, an interval of 9 seconds between the two. At a considerably higher intensity the first disappearance occurred at 116 seconds (A.D. 27, in an average of 4 trials) and the second at 153 (A.D. 27), an interval of 37 seconds. Four minutes complete dark adaptation was allowed between each trial.

The intensities of the lights employed in these experiments were not measured but the significance of the results is, nevertheless, obvious. The much longer time which is required for the brighter as compared with the duller light to show its first disappearance must be referred to the longer period needed for adaptation to bring the corresponding sensation near enough to the threshold for disappearance to be possible. Once this lowering of the potential of the visual cells has been accomplished, however, a second disappearance can occur without the lapse of an equal period.

After visual equilibrium has been reached, the frequency at which the disappearances take place will be determined by secondary causes, the most important of which will be considered shortly. If the equilibrium sensation is not a

[1] Ferree, C. E., 'The Fluctuation of Liminal Visual Stimuli of Point Area,' *Amer. J. of Psychol.*, 1913, 24, 378–409.

neutral mid-gray, and if its brightness is in some sense proportional to the intensity of the stimulus, the disappearances might be expected to be more frequent for the lower than for the higher intensities. With the 2° white stimulus at an intensity still higher than in either of the cases considered above, subjects *H* and *E* failed to observe any disappearances during continuous fixation periods of 30 minutes. Similar results were obtained with subject *E* when the spot was viewed for 30 minutes through a ruby red glass. At higher intensities the sensation from this latter stimulus was readily reduced to a pure yellow. A large number of experiments were carried out with a very bright red stimulus—containing only the orange and red of the spectrum—of visual angle approximating 6.4°, in which the fixation periods were varied from 7 to 10 minutes. Although the subjects (*E*, *H*, and *T*) reported apparent sensory equilibrium within 30 seconds to 4 minutes no cases of disappearance were observed. With a somewhat weaker light there was no disappearance even of the *R* component of the sensation. Similar results were obtained with a green stimulus of the same dimensions, except that no instance of complete disappearance of the green component was reported in the few trials made with this color.

VI. The Influence of the Size of the Pupil on the Visibility of Stimuli

The theory of Hering would explain the fluctuative disappearances of spot stimuli by complete reduction, through adaptation, of the original sensation to a neutral mid-gray. The fluctuations would then be regarded as *reappearances* due to eye-movement[1] or other external causes, which permit momentary phases of reversal of the adaptation (recovery). However, the force of this explanation obviously is gone if the conception which we have been attempting to support in the foregoing general argument be admitted, *viz.*, that equilibrium adaptation does not bring all sensations to a common level regardless of the nature of the stimuli which

[1] Ferree, C. E., 'An Experimental Examination of the Phenomena Usually Attributed to Fluctuation of Attention,' *Amer. J. of Psychol.*, 1906, 17, p. 83 ff.

cause them. The Hering explanation holds for sensations whose equilibrium levels place them beyond the possibility of discrimination with respect to the idioretinal gray, but not for those possessing a higher equilibrium than this. The question remains to be answered therefore as to whether fluctuations actually occur with such stimuli as the latter.

On the not rigidly critical plane of argument upon which we now stand, a method for the solution of this problem immediately suggests itself. As already noted, Ferree has shown that large angle stimuli of low intensity disappear and do not return.[1] This is due to the fact that eye-movements are too small to free the retina from continued stimulation, so that no opportunity is offered for 'recovery.' Such stimuli condition an equilibrium sensation which lies below the threshold. It is clear, then, that if a stimulus of large angular area disappears and then returns, while fixation remains as nearly as possible stationary, its disappearance cannot be a simple matter of 'fatigue' and 'recovery,' on any basis yet considered. If eye-movement is the only ocular change which can permit 'recovery,' any stimulus of large area which returns to visibility after a disappearance must have a *supraliminal equilibrium sensation.*

Experiment VI. The Fluctuation in Visibility of Large Colored Paper Stimulus Fields

That such large stimuli actually do show fluctuations can easily be proven. A half sheet of Hering red paper was arranged so as to give a stimulus of visual angle 35° by 24°, a dark fixation circle being provided in the middle. This was illuminated so as to have an intensity of three candles per square meter (as measured by the flicker photometer, subject *T*). With steady fixation, repeated fluctuations in the brightness of the field, with frequent, although momentary, complete disappearances, were observed by five subjects (*T*, *D*, *B*, *S*, and *M*) to whom the stimulus was presented. After heavy darkenings the field reappeared with even augmented brightness, and during fixation periods of 10 minutes and

[1] *Ibid.*, p. 85.

more, no *permanent* disappearance could be obtained, the actually observed obscurations being very brief. The stimulation was monocular, the other eye being heavily blindfolded.

Fluctuations of this sort obviously cannot be explained on the basis of complete fatigue followed by recovery due to eye-movement, because, under the conditions, deviations in fixation would not move the image on the retina through much more than one fiftieth of its total diameter. Changes in accommodation, also, could exert no influence on the course of the sensation, since with large uniform stimulus surfaces they could produce only edge effects upon the image. There exists, however, a very potent factor in the limitation of the action of an external stimulus upon the retina which, for some unexplained reason, has received comparatively little attention in general visual experimentation, namely, *the size and activity of the pupil.*

The diameter of the average human pupil is capable of variation from a maximum of 7.5 mm. to a minimum of 1.5 mm., although the usual aperture lies between 2 and 6 mm. Since the amount of light entering the pupil varies as the square of its diameter the intensity of a retinal image which, with a maximal pupil is 1, will with a minimal pupil be 1/25, *i.e.*, with the former an external stimulus intensity of one candle per square meter is visually equivalent to a stimulus intensity of twenty-five candles per square meter with the latter. Now although the size of the pupil is for the practical purposes of seeing a function of the external illumination, this size is by no means invariant during steady stimulation, oscillations as great as 2 mm. being quite common. All changes in the size of the pupil which occur in the presence of such stimulation must of necessity have an effect upon vision which is identical with that to be produced by actual variation in the stimulus intensity. That such variations are ideally fitted to cause fluctuative disappearances of the type we have just been discussing, a moment's consideration will suffice to show.

In his earliest writings Hering[1] pointed out that the dimming of a fixated stimulus field, which had been acting upon the retina for some time previously, might not only cause the stimulus to disappear visually, but, if the reduction in intensity were sufficiently great, might bring about a replacement of the original quality by its antagonist. This phenomenon is a well authenticated one, and has been studied in considerable detail by the present writer for certain special cases (which will be discussed in a subsequent article). For purposes of discussion it may be referred to as the *dimming effect;* and it clearly represents one aspect of the generation of the negative and complementary after-image. It is clear that this effect can follow as readily from a contraction of the pupil as from the turning down of a lamp or the closing of a shutter.

Experiment VII. The Correlation of Movements of the Iris With Visual Fluctuations

In order to determine empirically whether the fluctuations occurring in the visibility of stimuli supposed to have a supraliminal equilibrium sensation can actually be attributed to pupillary contractions and expansions, the following experiments were made. The stimulus surface last described above was perforated at the fixation spot, and a reading telescope was placed behind it in line with the perforation and the eye of the subject, who was provided with a signal key. In this way the subject's pupil could be observed at the instant when disappearances of the stimulus were reported by a depression of the key. Observations on five subjects (*T*, *M*, *B*, *D*, and *S*) were made with this arrangement, the iris being illuminated no more than necessary to permit its movements to be clearly seen through the telescope.

The results of these experiments leave no doubt that pupillary contractions and expansions are a primary cause of the fluctuation of stimuli, which have a supraliminal sensory equilibrium. In every series of observations striking coin-

[1] Hering, E., 'Zur Lehre vom Lichtsinne,' *Sitzb. d. Wien. Akad.*, 1872–1874, p. 132.

cidences between pupillary contractions and the disappearance signals were noted. Practically every strong contraction was followed immediately by a signal. The best conditions for observation appeared to exist when the activity of the pupil became rhythmic. As a sample of the results obtained, the following trial with subject *D* may be described. At the beginning of the fixation the pupil had a fairly large opening, about 4 to 5 mm., and showed some fluctuations rather difficult to follow. During the preliminary period the subject gave some disappearance signals for which a clear correlation with pupillary action could not be affirmed. After the lapse of an interval which would seem to correspond roughly with the time required for the retina to reach sensory equilibrium, the pupil began rather suddenly to show sharp rhythmic contractions, in which the diameter changed from about 4 to about 2 mm. During a continuous series of twenty of these contractions there was only one for which the subject did not give a signal at an interval of 2/5 to 3/5 of a second after its completion, a period which corresponds very closely to that required to produce the dimming effect. The subject reported each time the (Hering red) color wholly disappeared from the field of vision. After a rest the observations were resumed, with exactly similar results in every respect. These two series then give 38 cases of perfect coincidence between pupillary contraction and disappearance of the stimulus, out of 40 cases of observed contraction, and during the periods of easily observable contractions there were no cases of disappearance which were unaccompanied by the pupillary change. The subject stated that the disappearances which occurred in the preliminary periods of small fluctuations were not as well marked as those which followed.

There appears to be evidence in these observations of a definite rhythmic reflex of the pupil, the biological significance of which may perhaps lie in the possibility of producing clear intermittent vision where satisfactory continuous vision is impossible owing to adaptation. The contraction of the pupil will have an effect similar to that attributed to eye movement in the case of small luminous stimuli, and permit the retina to recover, so that in the succeeding phase of dilation vision is again restored. Such a mechanism would of course apply primarily to stimuli of subliminal sensory equilibrium although it would not be limited to these. The assumption of its existence is not necessary to validate the general argument regarding fluctuations in visibility which we are here presenting, since wavering aperiodic activity of the iris is an unquestionable fact.

An interesting variation of the above experiment is the following. The experimenter informed the subject that he would tell him how the stimulus field looked to him (the subject) from time to time. This the experimenter did by inference from the pupillary size at the moment. Thus, when the pupil was large he inferred an apparent brightness of the stimulus, with contraction he stated that it was now darkening, etc. At the end of a three-minute experiment of this sort the subject (*B*) reported that not more than three mistakes had been made in the 'predictions' of the experimenter.

In these experiments the writer several times exchanged places with his subject, with a confirmation of the above described results in all essential particulars. Limitations of time did not permit the application of a more objective method for recording the pupillary changes contemporaneously with the disappearance signals, a development of the research which is of course desirable.

The correlation between disappearance and obvious pupillary contraction is not perfect. When the activity of the iris was of the wavering type a considerable number of signals were given by the subject for which the experimenter noted no distinct contractions, and on the other hand, there were many contractions for which no signals were given. The latter point is not of great importance, since whether or not a contraction produces a disappearance will depend both upon its magnitude and upon the exact state of adaptation at the moment, and it must be kept in mind that with an active pupil this state of adaptation is by no means constant even in the presence of a constant external stimulus. With regard to the report of disappearances when no contractions were observed, it may be said in the first place that, except where the changes in pupillary size have a considerable amplitude, they are not easy to record by the method of observation here employed, yet states of adaptation may occur where a small contraction will suffice to cause a disappearance. Contractions which start with a wide pupil, if themselves small, do not alter the intensity of the retinal stimulus by a large fraction, but they may nevertheless be effective in producing disappearances, owing to the fact that they act upon the basis of a strong prior 'fatigue.' On the other hand, linearly small contractions which start with a narrow pupil cause a relatively large change in the illumination of the retina.

It is a well-known fact that, in the so-called accommodation reflex, the pupil undergoes contraction simultaneously with the adjustment of the lens of the eye for near vision. Consequently the observations reported by other investigators which go to show that changes in accommodation coincide with phases of disappearance of visual stimuli, may receive an interpretation in terms of the activity of the iris. For stimuli of very small area, any accommodation change which causes blurring may produce a dimming effect by spreading the light out over a larger area. However, the probable efficacy of this change in the production of disappearances is doubtful, on account of the fact that this same process of spreading brings the light to fresh retinal surfaces, and it is a well-known fact that for small areas the threshold depends upon the total light energy received rather than upon the illumination. The effect would be limited in any case to very small stimuli.

In their crude outlines the above reported experiments do not prove that the stimuli which were observed to fluctuate were actually of supraliminal sensory equilibrium, since it is obvious that a stimulus which disappears through simple fatigue may be brought once more to visibility through a dilation of the pupil—the effect of which upon the retina is the same as brightening the stimulus. This will be true especially if the dilation is preceded by a contraction. However, the fact that, in the experiments described, the subjects' signals practically always followed the pupillary contractions by a considerable interval rather than *vice versa*, would seem to indicate that the disappearances were caused by the contractions and not the contractions by the disappearances. If this is the case the disappearances must have been of the dimming effect type and the stimuli, consequently, of supra-liminal sensory equilibrium. Another fact which supports this latter interpretation of the results lies in the momentary character of the disappearances, which was paralleled by the almost instantaneous redilation of the pupil. Long periods of visibility compared with those of invisibility would appear to indicate that the former represent the equilibrium state of the organ. As will be shown later, the effective intensity of the 3 candles-per-square-meter stimulus used in the above experiments is about ten times that of the minimal stimulus which, with the artificial pupil, shows no disappearances.

It is obvious that with stimuli of high intensity, for which the pupil has attained its maximum contraction, a pupillary dimming effect is impossible. This consideration harmonizes with the fact, which was pointed out in section III, that disappearances of stimuli of high intensity do not occur.

VII. Observations and Measurements Upon Equilibrium Sensations With the Use of the Artificial Pupil

The crucial test of the theory that the primary cause of the fluctuations in visibility for other than minimal stimuli is to be found in pupillary contraction and expansion would seem to lie in the elimination of these changes either by atropinization or by the use of an artificial pupil. Experiments have already been carried out by Pace[1] and McDougall[2] in which the eye muscles were atropinized, although the primary object of the treatment was apparently to paralyze the ciliary muscles rather than those of the iris. Both of

[1] Pace, E., 'Zur Frage der Schwankungen der Aufmerksamkeit nach Versuchen mit der Masson's Scheibe,' *Phil. Stud.*, 1893, 8, 388–403.

[2] McDougall, W., 'The Physiological Factors of the Attention Process,' *Mind*, 1903, 26, p. 476.

these observers reported fluctuations. However, their methods of work were not such as to yield conclusive results with regard to that sort of influence of the pupil in which we are now interested, McDougall's primary interest being in retinal rivalry phenomena. Consequently further work along these lines would seem to be desirable. The employment of the artificial pupil provides a better means for the elimination of pupillary effects from the adaptation process in vision than do the drugs, atropin and pilocarpin. Such a pupil, of diameter 2.36 mm., was used by the writer in all of the measurements to be described below. In the course of these measurements several thousand observations of spot stimuli of medium brightness—ranging in size from .7° to 3.5°—on a dark ground, with fixation periods varying from one to 15 minutes, were made, and the number of fluctuative disappearances which were seen could be counted on the fingers of one hand. This would appear to be conclusive proof that the variations in size of the natural pupil are the essential cause of these disappearances, when the pupil is free to modify the intensity of the retinal stimulus.

The above statement applies only to stimuli of moderate or of high intensity, and not to those of very low intensity. If the latter are what we have designated as 'of subliminal sensory equilibrium' it is obvious that adaptation will cause their disappearance, and that when their equilibrium sensations lie very close to the threshold,[1] slight disturbances, such as slips of fixation, changes in the nutrition of the retina, etc., will cause them to fluctuate in visibility.

Experiment VIII. Measurement of the Minimal Stimulus Intensity Necessary to Prevent Fluctuative Disappearances with the Artificial Pupil

In order to determine the intensity of such an equilibrium threshold stimulus the following experiments were made. The results apply, of course, only to the special conditions under which the measurements were taken, *viz.*, for foveal vision with a dark contrast field.

[1] The threshold concept throughout this paper is clearly differential as well as absolute.

The stimulus employed in each case was in the form of a circular spot 83 cms. distant from the eye of the subject and of angular diameter 1.28°, which, on the basis of the estimates usually given, could be counted upon to stimulate only retinal cones, in the normal subject. In the case of the blue and the green, where signs of a photochromatic interval appeared, a field of size .69° was also employed. In the center of the field was placed a dark circular fixation spot of angular diameter .18°. The lights employed in these fields consisted of radiation selected and controlled by a spectrometer system. The four wave-length groups, used, corresponded approximately to the subjects' four psychological primaries, *R*, *Y*, *G*, and *B*, and the flicker equations upon which the candles-per-square-meter values in Table III. are based were made by the same subjects (*S* and *T*).

These observations and all subsequent ones in which spot stimuli were employed were carried out in a room which, as regards the vision of colors, could be considered practically dark. What little diffuse light there was present was prevented from affecting the stimulus field or acting on any but the extreme peripheral parts of the retina during the observations. The periods of rest given to each eye after an experiment before it was used again were adequate to insure a return to maximal color sensitivity.

In the present experiment the intensity of the spot stimuli was under the control of the subject, who was instructed to fixate the black mid-point and, starting with a very low intensity, to increase the brightness slightly every time a disappearance occurred. The time of each disappearance was noted, and when none was reported during a three-minute interval the stimulus was regarded as of the required threshold intensity. At least four successive determinations of this sort were made upon each color. The corresponding voltage readings on the spectrometer lamp were then averaged and the intensity of the field corresponding to this average was photometered. The final results are given in Table III.

As shown in the table it was easy to find a point at a low intensity for which no disappearances of the spot occurred.

In the case of the red and yellow the stimulus retained its original color, in good saturation, at the intensity in question. With the green and blue, however, there was a strong tendency for it to pass over into a gray before disappearing. This well-known phenomenon of a photochromatic interval in the case of lights of the shorter wave-lengths, made it necessary to modify the experimental method for stimuli of this sort so that the threshold intensity value finally found was one for which the *color* remained constantly visible, although a considerably lower value might have sufficed to keep the spot in view as a gray. The photochromatic interval is usually supposed to be absent with strictly foveal stimuli,[1] and the few observations here reported cannot be regarded as adequate to refute this view, since slight slips of fixation easily bring the stimulus into the extra foveal region when it has fallen below the threshold in the fovea itself.

The observations and measurements just described show that if the action of the iris is eliminated, a stimulus of low intensity is able to maintain a steady sensation without loss of its proper quality for long intervals. The periods of time employed are probably sufficiently protracted to justify the assumption that the measurements apply to the equilibrium sensations of the stimuli which were used. These results, then, argue very strongly against the Hering view that equilibrium sensations are a uniform neutral gray, and in combination with facts previously considered constitute proof of the efficacy of the natural pupil as a cause of visual fluctuations.

The fluctuations which are actually observed near the threshold are probably somewhat complex in their causation. Some of them are no doubt to be explained on the eye-movement-recovery theory, either in its ordinary form or with respect to the influence of these movements on the amount of light entering the eye through the artificial pupil. Large deviations from 'register' (co-centering) of the natural and artificial pupils cut off some of the light and this produces a dimming effect. Registration was secured at the beginning

[1] Nagel, W., 'Adaptation, Dämmerungssehen und Duplizitätstheorie,' *Helmholtz's Physiol. Opt.*, 3 Aufl., 1911, 2, p. 309.

Table III

The Value of the Equilibrium Threshold Stimulus for Four Spectral Lights, with an Artificial Pupil

General Conditions: Size of stimulus field: 1.28° in diameter (circular), dark fixation point: .18° in diameter; Total time of each fixation: about 15 minutes; Standard time during which the spot must show no disappearances of color in order that the intensity should be taken as a threshold: 3 minutes; Diameter of artificial pupil: 2.36 mm.; Fixation, monocular, the other eye being blindfolded.

Explanation of the Table: In the first column, below, are given the wave-length ranges employed, in $m\mu$. The second gives the thresholds in candles per square meter for the two subjects *S* and *T*. The numbers in parenthesis represent the intensity of the stimulus in photons or the product of the candles-per-square-meter value and the millimeter area of the artificial pupil. The threshold values themselves are given under "Av." and are the average of at least four separate determinations. Under "A.D." is given the estimated average deviation of the mean, independent of photometric errors, which may amount to 2 percent, and the 5 percent uncertainty involved in the reproduction of the Hefner amyl acetate standard.

Radiation Range, $m\mu$	Threshold Intensity in C. per Sq. M.				Color
	Subj. *S.*		Subj. *T.*		
	Av.	A.D.	Av.	A.D.	
690.2–705.8	.436 (1.90)	.021 (.09)	1.088 (4.75)	.045 (.20)	*R*
580.0–587.6	.85 (3.71)	.54 (2.4)	.94 (4.10)	.13 (.57)	*Y*
525.4–530.0	1.782 (7.77)	.018 (.08)	.94 (4.10)	.13 (.57)	*G*
451.2–455.8	2.09 (9.11)	.12 (.52)	1.50 (6.54)	.23 (1.0)	*B*
General average	1.29 (5.63)		1.12 (4.78)		

and tested at the end of each trial in the above series by a method described elsewhere.[1] So far as the writer's own observations are concerned, however, it appears to him probable that the fluctuations of liminal color fields should be attributed to that sort of disturbance in the nutrition or electrical condition of the visual system which goes with the undulations in the idio-retinal light. Ferree's[2] observations in this respect

[1] Troland, L. T., 'The Theory and Practice of the Artificial Pupil,' *Psychol. Rev.* 1915, 22, 172–176.

[2] Ferree, C. E., 'The Streaming Phenomenon,' *Amer. J. of Psychol.*, 1908, 19, 484–503.

seem to him correct, since by far the greater number of disappearances of such color fields involve a submergence of the spot in one of these bright surgings. Another cause is probably binocular rivalry with the opposite wholly dark field. It is clear that the above measurements of the equilibrium threshold, which indicate an average value of about 1.2 candles per square meter with a pupil of 2.36 mm. are not minimal, because they indicate a factor over and above that required by considerations of pure adaptation, which is necessary to counterbalance the disturbing influences just mentioned. Moreover, it would be absurd to claim that, with prolonged fixation, cases might not arise in which this factor would have to be increased considerably.

Experiment IX. The Equilibrium Sensation for Spot Spectral Stimuli of Equal Brightness, with the Artificial Pupil

The experiments just described were extended in a series of observations made by subject *B* in which spectral stimuli of angular size 1.28°, only, were used, with fixation conditions similar to those in the first series. Eight lights were employed which were equated in respect of brightness by means of the flicker photometer, the common intensity being 44.7 candles per square meter, an artificial pupil of 2.36 mm. being utilized, yielding a photon value of 195.1. The ranges of wavelengths employed were the four used above, together with a violet, blue-green, yellow-green, and orange. The flicker equations were made by subject *L* for another purpose, and were not verified for subject *B*, because of limitations of time and the fact that both were thoroughly normal in their color vision, subject *L* having himself done extensive work upon color. The results of the ten-minute monocular fixation of these stimuli are given in Table IV.

The primary object of this series of observations was to determine whether spot stimuli of moderate intensity show radical changes in apparent quality at sensory equilibrium. The results indicate in the first place that with the artificial pupil no disappearances of such stimuli occur, and that even

Table IV

The Fatigue Changes in Sensation with Small Spectral Stimuli of Equal Moderate Brightness, with the Artificial Pupil

General Conditions: Size of stimulus fields: 1.28° in diameter (circular), dark fixation point: .18° in diameter; Total time of each fixation: 10 minutes; Diameter of artificial pupil: 2.36 mm.; fixation: monocular, in the dark; Common intensity of the stimuli: 44.7 candles per square meter, or 195.1 photons. The precision of these and subsequent photometric measures is about 2 percent as regards the standard (the Hefner, Reichsanstalt, amyl acetate lamp) used by the writer, and about 5 percent as regards the general reproducibility of this standard. The precision of reproducibility of the stimulus fields varies from 0.5 percent to about 2 percent. The precision measures are expressed in terms of candles per square meter. Subject: *B*.

Explanation of the Table: The arrangement and symbolism in this table are similar to those in Table I., *q.v.* The first column gives the wave-length range of the stimulus.

Radiation Range and Color	Color Phases			Remarks
	Initial	Intermed.	Final	
R 690.2–705.8	*r′y″w*	*r″w*	*r′w*	
RY 605.6–615.0	*r″y*	*r′w*	*r′y*	
Y 580.0–587.6	*y″w*	*y′w*	*yw′*	Strong return of yellow.
YG 552.0–558.0	*y*	*g″y*	*g″y*	Green wholly gone at first.
G 525.4–530.0	*yg*	*g′y*	*y′g*	
BG 487.4–494.4	*g″bw*	*g″bw*	*g″bw*	
B 451.2–455.8	*bw*	*bw*	*bw*	Not much change.
Vi 407.2–417.6	*b′w*	*b′w*	*b′w*	Red component permanently disappears.

very long fixation periods do not entail a reduction of the sensory quality to median or any other sort of gray. They indicate, moreover, that in the case of the psychological primaries there is no loss of the essential hue. The intermediate tones, however, may easily lose the *R* and *G* components, especially when the latter are relatively weak at the outset. The violet entirely lost its *R*, which did not return. The orange, which at the start was strongly reddish, showed marked diminution of this quality, but never lost it entirely. The blue-green always retained a slight tinge of green, and the yellow-green, while at the end of one minute it became pure yellow, during the last five minutes of the fixation showed a perfectly clear greenish tone. One of the most interesting results of the present series, therefore, is the confirmation which it provides of the observations made by subject *T* on adaptation to large fields of stimulation, *i.e.*, that not only is the maximum depression of the sensation reached quickly (in one to three minutes), but that after this time there is a revival of the original quality.

During the course of the work with the method of critical flicker frequency to be described in a second paper, a considerable number of long interval fixations of the same spectral colors used above were carried out. The light in these cases was interrupted at the rate of 25 alternations, of equal light and dark periods, per second, the brightness of the stimulus being, during the greater part of the time, very close to that required to eliminate flicker at this frequency, approximately 30 candles per square meter. The fixation periods were nine minutes or more, subject *T* being the observer. The dimensions of the stimulus spot were the same as specified for the series just described. At least four cases with each color, two for each eye, were recorded. The results were that the only stimuli completely to lose any of their characteristic components were the violet and the blue-green, which lost their *R* and *G*, respectively, both becoming a whitish, unsaturated blue. All of the colors of course became less saturated and darker. The *G* component did not disappear from the yellow-green or the *R* from the orange. The red became slightly bluish instead of yellowish. The longest cases of fixation were of the red for 30 minutes, of the green for 21, and of the blue for 15. It is worth noting that practically no after-images were visible at the end of these long fixation periods, when the stimulus was removed and the eye left in the dark. This latter observation was repeatedly verified by subject *H*.

Experiments were not made with spot stimuli of high intensity, since the phenomena which they exhibit are quite well known. 'Fatigue' and increased intensity coöperate to extinguish, first, the *R* and *G* components in the sensation and then all chromatic elements, so that the stimulus appears white. This occurs with pure spectral lights at sufficiently high intensity, as well as with heterogeneous chromatic stimuli.

Hering's theory of color adaptation is inconsistent with the notion that a visual quality once below the threshold should return without alteration in the intensity of the stimulus. This statement would hold also of the doctrine which the present writer favors, provided the stimulus in question is the one under the influence of which the sensation originally disappeared; that is, if this stimulus is one of subliminal sensory equilibrium. But upon Hering's assumptions all stimuli are of this character.

Suppose that the retina is stimulated for some time, less than that necessary to produce equilibrium, by a light of relatively high intensity. This, in the words of Hering, will cause a lowering of the 'potential' of the visual cells. If, now, the light is dimmed, the sensation will easily fall below the threshold and may even be replaced by its antagonist. Under these conditions, if Hering's hypotheses are valid, the sensation cannot again mount into consciousness, since either the sense-cell is already at equilibrium with the stimulus of lowered intensity or the equilibrium condition will be approached from the side of the antagonist of the original sensory quality.

Experiment X. The Reappearance of a Visual Quality, Lost Through Adaptation, under Continued Stimulation

To test this issue, experiments were made with the four primary spot spectral stimuli used in Experiment VIII., the usual artificial pupil being employed. The greater part of the work was done by subject *S*, but his observations were corroborated by subject *T*.

The eye was first exposed for 20 seconds to a stimulus of about 30 candles-per-square-meter intensity (artificial pupil: 2.36 mm.), followed by a rapid decrease in the intensity, until the subject signaled that the spot had disappeared. Fixation was kept as nearly as possible constant by means of the bright crescents appearing at the edges of the adapted area and a signal was given when the whole area of the field came again into view.

The point to which it was necessary to reduce the intensity of the stimulus before disappearance occurred, depended upon the original brightness and the time of preëxposure. In all cases the spot reappeared almost immediately, and remained visible thereafter as long as fixation was continued. In the case of the blue the reappearance was often that of a gray rather than a blue spot. The quantitative measurements made upon these phenomena were insufficient to permit any very secure generalizations, but they strongly suggest the view that regardless of the degree to which a specific sensitivity of the visual cells has been reduced by prior stimulation, this specific sensitivity will automatically readjust itself to a succeeding stimulation, with an end result which is practically independent of the preliminary 'fatigue.' The experiment proves, in general, that Hering's doctrine of the neutral median character of all equilibrium sensations cannot be true, since in these observations the movement towards equilibrium may be away from the neutral mid-gray and in the direction of the quality proper to the stimulus.

Objections to the above conclusions, based on the notion that recovery here may have taken place during an eye-movement incident upon the subjective disappearance of the stimulus spot, cannot be admitted, since lapses of fixation were easily guarded against by means of the bright crescents which they yielded on the edges of the fatigued area. The process of recrudescence of the sensation could be clearly observed to occur within the border defined by these crescents. With care, a practiced observer can maintain his fixation of a circular spot of the size here employed, to within about one sixteenth of its diameter.

VIII. The Influence of Binocular Rivalry upon the Visibility of Monocularly Viewed Stimuli

The work with the artificial pupil which has been described above shows that the liability of specific visual qualities to suffer extinction through processes of adaptation is much less than is commonly supposed. The experiments indicate very clearly the preëminent importance of pupillary fluctuations in bringing about the much-discussed fluctuative disappearances of visual areas. However, it would be folly to claim that the contractions and expansions of the iris are the only changes which can aid such disappearances. Certainly other factors are operative, and among these, although its conditions of operation are somewhat special, is the process underlying retinal or binocular rivalry. A large number of very interesting observations by McDougall[1] reveal the importance of this mechanism in causing visual fluctuations, and give clear indications that it is not without effect even upon supposedly monocular vision. In the majority of exact studies of visual sensation it is necessary to stimulate only one eye, the other being in the meantime blindfolded. From a formal point of view this situation appears favorable for rivalry phenomena, since the two fields of vision are qualitatively different. However the dark gray of the blindfolded eye appears at most to have but a weak tendency, if any, to oust the quality of the opposite visual field from consciousness. As shown by the following experiments there are conditions under which this impotency of the non-externally stimulated field disappears.

Experiment XI. The Binocular Interaction of an Adaptation After-effect with the Original Stimulus

The following observations were made by subject *B* under the conditions specified in Experiment IX. After a long preëxposure of one eye to the spectral spot stimuli described above, it was found that if the stimulus was immediately transferred to the opposite eye it at first appeared of the proper color, but as soon as fixation was established it became a sparkling white, which gave place rather gradually to the proper quality. This effect was obtained several times each with the *R*, *Y*, *G*, *YG*, and *BG*, but not with the *B* or *Vi* stimuli. In the case of the orange a similar phenomenon was observed, in which, however, the crystalline appearance was tinged with yellow. With the green the whiteness changed after the lapse of a few seconds to a pure yellow, which persisted for a long time. In the other cases it was replaced after about 20 seconds by the normal stimulus

[1] McDougall, W., 'Some New Observations in Support of Thomas Young's Theory of Light and Color Vision,' *Mind*, 1901, 26, 52–98, 210–245, 347–382.

quality. The white was described as very brilliant and possessed of the characteristics common to binocular luster, so that the idea is clearly suggested that it was the result of a central combination of the after-image process of one eye with the primary image of the other, to form a complementary white. Time did not permit a careful investigation of this phenomenon to determine whether or not it could be obtained with blue and violet under the right conditions.

In the course of the work with large colored stimuli described in Experiments I. and II., it was found by subject *T* that if, after prolonged fixation with the right eye, the left eye was exposed to the stimulus and the right blindfolded, strong periodic darkenings, characterized by a pebbly pattern and a tinge of the hue complementary to the stimulus, appeared in the field. These darkenings were very patchy—and seemed generally to omit the foveal region—but where they occurred they entirely obscured the stimulus. If after a short fixation period with the left eye, the observation was continued once more with the right, similar surgings of the field were observed, except that in this case they were usually not darkenings, but consisted of strong patches of a color different from that of the stimulus. As in the case of the left eye the patches entirely supplanted the proper quality of the stimulus, except when—following the procedure of Experiment III.—a neutral comparison object was introduced, which permitted the original quality of the stimulus to be perceived through a mist of the opposing quality. For the orange, yellow, yellow-green, and blue-green papers this latter quality was purple. The red paper showed bright yellow patches, the blue simply darkenings, the purple bright green blotches, and the violet a general whitening of the field. In the case of the green, the purple color persisted during two minutes' continuous fixation in sunlight. In all of these observations only one eye was externally stimulated at one time.

It would be foreign to the purpose of the present paper to attempt a theory of the phenomena just described. Their sufficient significance in the present connection lies in the proof which they offer of the fact that, under certain conditions, the disappearance of the primary quality of a stimulus may be due to processes initiated in the retina of the eye opposite to that exposed to the action of the stimulus in question.

IX. Conclusion

The purpose of the present paper has been to show that—contrary to the requirements of the Hering theory in its ordinary form—a visual sensation is not a process—like those of the classical theory of available energy—which goes on only at the cost of its final self-destruction. That specific adaptation reduces the sensitivity of the visual cells to a very marked degree cannot be gainsaid; but that this reduction is such as, with continued stimulation, to bring all stimuli to equivalency it is the object of the foregoing admittedly incomplete studies to disprove. Stimuli of low intensity may readily be carried below the threshold by adaptation, but the tests and measurements described in the present paper demonstrate that by far the greater number of visual stimuli which act upon our eyes in everyday life are of the constantly supraliminal equilibrium type both with respect to brilliance and to saturation.

In another place the author[1] has suggested a modification of the Hering theory which will permit one to account for the elevation of the threshold through adaptation, and which at the same time sets definite limits to this process, which are determined by the nature of the stimulus. This view provides a basis for the theoretical study of sub- and supraliminal equilibrium sensations in general. The essential conceptions and assumptions involved may be outlined briefly as follows.

Let us suppose that the intensity of a given component, V, of the visual sensation is such a function of a katabolic rate of change of the concentration, s, of the visual substance, S, in the retinal end cells as we have previously defined. Chemical principles lead us to suppose that this katabolic rate, q, will not be zero even in the absence of light, and that

[1] Troland, L. T., 'Adaptation and the Chemical Theory of Sensory Response,' *Amer. J. of Psychol.*, 1914, 25, 500–528.

its magnitude will be proportional at all times to the concentration of the substance at the time. If we assume the reaction to be irreversible, the spontaneous decomposition in question may be represented by the equation:

$$(10) \qquad k = c_k s,$$

where c_k is a constant. Similarly, the component of the rate due to the action of light may be represented by the equation:

$$(11) \qquad d = c_d i s$$

in which i is the light intensity. In order to maintain the process, repair is necessary, and if this repair consists of a flow of substance into the cell which is governed by osmotic principles, it should follow an equation similar to:

$$(12) \qquad a = r - c_r s,$$

where a is the rate of inflow, c_r a constant, and r another constant, whose magnitude depends upon the concentration of the nutritive substance in the lymph outside the cell.

It will be clear from a consideration of the above equations that, on the basis which they provide, the total rate of change of s, which latter variable is equivalent to the sensitivity of the cell, must be

$$(13) \qquad ds/dt = a - k - d$$

or

$$(13a) \qquad ds/dt = r - (c_r + c_k + c_d i)s.$$

But at equilibrium $ds/dt = 0$, for which value equation (13a) can be reduced to:

$$(14) \qquad s = r/(c_r + c_k + c_d i),$$

which states that the more intensive the stimulus with which the visual cell comes into equilibrium, the lower will be its equilibrium sensitivity. Now, we have assumed the intensity of the sensation to be proportional, in some sense, to the katabolic component of the change in s. This component obviously has the value:

$$(15) \qquad q = (c_k + c_d i)s,$$

or, for equilibrium, from (14):

$$(16) \qquad q = r(c_k + c_d i)/(c_r + c_k + c_d i),$$

in accordance with which the rate of the katabolic process at equilibrium must be higher the higher the intensity, i, of the stimulus.

If, in equation (16), we place $i = \infty$, we find $q = r$, which shows that, with increasing stimulus intensity, q approaches r as a limit, thus revealing equation (16) as a statement, for equilibrium conditions, of the general relationship formulated in Fechner's law. An analysis of the value of di/dq yielded by the equation in question shows that it accounts for the so-called upper and lower deviations from Fechner's law, as well as for the general relation. If we place $i = 0$, we find from (16) that

$$q = rc_k/(c_r + c_k), \tag{17}$$

which corresponds with the sensation from a retina at equilibrium with darkness, or with rest idio-retinal light. If we suppose the discrimination threshold for the visual component, v, stated in terms of q to be Δq, it is obvious that a given stimulus will be of the supraliminal equilibrium type if its intensity, i_p, is such that:

$$\frac{r(c_k + c_d i_p)}{(c_r + c_k + c_d i)} - \frac{rc_k}{(c_r c_k)} > \Delta q \tag{18}$$

and will be of the subliminal equilibrium type if the inequality is reversed in direction.

In considering the question as to the causes underlying the fluctuations of visual sensation, one must carefully avoid preconceptions, since this field of investigation is extremely full of complexities. The view that such fluctuations are due to deviations of the attention seems obviously to have been supported by prejudice rather than by either theoretical or empirical analysis of the facts. However the writer deems it probable as a result of certain observations of his own that, especially with threshold stimuli in the dark, the whole visual field—so far as regards detail—may sometimes lapse from consciousness, by a process seemingly describable as failure of attention. That attention to visual objects goes regularly in the form of waves, which set a definite limit to the length of time during which such an object can be continuously perceived, seems to the writer, on the basis of the work described above, very doubtful.

Ferree[1] has attributed certain visual fluctuations to the effects of 'the streaming phenomenon,' or the surgings of the idio-retinal light. As already noted, the present writer's observations tend to bear out this view with regard to liminal stimuli seen with the artificial pupil. When the idio-retinal field is quiet there are far less marked fluctuations in the visibility of the liminal stimulus, and the intensity which is necessary in order to maintain it constantly in view is lower than when strong surgings of the idio-retinal light are present. The blotting out of spot stimuli by the idio-retinal surgings appears to be a matter rather of swamping than of inhibition. If the intensity of the stimulus is increased so that the sensation which it causes is always perceptibly greater than the brightest of the idio-retinal waves, it shows no disappearances. Accordingly it seems probable that disappearances due to the streaming phenomenon should be attributed to momentary increases in the value of the second complex term in equation (18), which causes the difference between the two terms to become less than Δq.

Momentary disappearances of a visual stimulus may sometimes be due to rapid eye-movements which, as Holt[2] has shown, are accompanied by central anaesthesia. Even if there were no such anaesthesia, the sensation might lapse, in the case of minimal stimuli at least, owing to the speed of the movement and the consequent lowering of the time of action of the stimulus upon a given retinal point. In such cases the stimulus would reappear again almost immediately, as fixation was reëstablished. Some of the observed disappearances of very dim lights are so nearly instantaneous as strongly to suggest an explanation of this sort.

The reappearance of stimuli of subliminal sensory equilibrium as a result of recovery during eye-movement is also,

[1] Ferree, C. E., 'The Streaming Phenomenon,' *Amer. J. of Psychol.*, 1908, 19, 484–503.

[2] Holt, E. B., 'Eye-Movement and Central Anæsthesia,' Psychol. Rev. Monogr. Suppl. No. 17. *Harvard Psychol. Stud.*, 1903, 1, 3–46.

of course, a distinct possibility. This would apply only for values of the stimulus intensity, i, for which the inequality of (18) has a reversed direction.

Phenomena of binocular interference, or retinal rivalry are probably of importance for the explanation of certain special cases of disappearance, as shown by the observations described in the just preceding Section.

When considering the problem of the quality of the equilibrium sensation with respect to Hering's views it is especially necessary to treat independently of brilliance and of hue. Given components of the original sensation may disappear at equilibrium without necessitating that the whole should fall to the level of mid-gray. Adaptation weakens all of the chromatic components more than it does the achromatic. Since it seems likely that each of the six fundamental components of the visual sensation has a retino-cerebral apparatus which is more or less independent of that of the others, each of them may be considered as having separate equilibrium conditions, which must be quantitatively determined before any definite prediction can be made concerning their behavior. The experiments which we have reported above indicate that a spectral stimulus, for example, may be of the subliminal equilibrium type with respect to one sensory component (say R), while it is not so for another (as B). All experiments which demonstrate a photochromatic interval at the threshold for equilibrium must be interpreted from the same point of view, the equilibrium conditions for the hues involved being considered independently of those for the 'gray' component.

It is necessary, however, with regard to the disappearance of part components of the visual sensation, to consider not only the independent properties of these components but also the laws of their mutual interaction, in particular their relations of antagonism or complementation, and the manner in which they are capable of overwhelming one another. The disappearance of R and G at high intensity equilibrium may be assigned not so much to the fact that they have low absolute equilibrium sensation intensities as that these intensities are *relatively* low compared with those of Y, B, and W.

It is probable, also, that at high intensities the relation of antagonism enters in, since if each of the components approaches an asymptotic limit with increasing intensity, the same values will be attained by antagonistic excitations at high intensities of any stimulus, regardless of the wave-lengths which it includes, and it is naturally to be assumed that at such limits the antagonists, being probably at about equal strength, will either inhibit each other or combine to form white. In either case the two antagonistic qualities will be eliminated as possibilities in sensation. For very low intensities it is probable that, through dark adaptation, a marked increase in the sensitivity of the achromatic mechanism, relative to that of the chromatic, may cause a disappearance of color components through 'swamping.' This is probably true in the case of the achromatic interval in the blue, since at both very low and very high intensities the blue of the spectrum yields a sensation having a strong W component.

All of the above considerations apply to observations with the artificial pupil as well as to those in which the natural pupil is free to modify the effective intensity of the stimulus. In the latter case, however, the writer is convinced that with stimuli of ordinary intensity, practically all of the observed fluctuative disappearances are due to the activity of the iris. Several types of fluctuation caused by the pupil may be suggested. The first applies to a subliminal equilibrium stimulus. We will suppose that the sensation has reached equilibrium with a fairly wide pupil and that it has fallen below the threshold. If the pupil now contracts, recovery will set in, that is, the level of sensitivity will be raised in accordance with equation (14), owing to the weakness of the effective stimulus with which the organ is in equilibrium. When the pupil expands once more the stimulus momentarily reappears. It is clear that whether a given external stimulus is one of subliminal or of supraliminal equilibrium may be determined by the size of the pupil itself, since this affects the energy reaching the retina just as does an actual variation in the external illumination. Hence a second case may be suggested in which a subliminal equilibrium is reached with a small pupil, this being followed by pupillary expansion such that the increased effective intensity now yields a supraliminal equilibrium sensation. In this case the stimulus should not disappear again until another contraction of the pupil ensues. A third type of fluctuation caused by the pupil would be that of a simple 'dimming effect,' which requires that the visual cells be fatigued by a light of relatively high intensity—wide pupil—which is then dimmed—narrowing of the pupil. A dimming effect of this sort may be produced regardless of whether the stimulus is one of sub- or of supraliminal equilibrium intensity, although with the former alternative the disappearance would be only momentary, even if the pupil remained contracted. Two cases are really involved here, however, one of the simple dimming effect, and one in which the pupillary contraction brings the effective stimulus below the equilibrium threshold. It is probable that the simple pupillary dimming effect is the principal cause of fluctuations in free vision, a view strongly supported by the fact of their absence with stimuli of high intensity.

The argument of the present paper, directed as it is against Hering's conclusion that all equilibrium sensations are neutral mid-gray, nevertheless admits that complete disappearances of visual stimuli—such as Hering adduces in support of his theory—actually occur. However, it is intended to prove that such disappearances take place only under definite special conditions and should not be interpreted as demonstrating the evanescent quality of all fixed visual sensations. What some of the special conditions are under which disappearances do occur, we have tried to make clear, as well as to show how these disappearances are consistent with a theory of visual response which does not, like that of Hering, imply a common dead level for all equilibrium sensations.

6. Mention has been made of the factor of expectation in the present problem. In Breitwieser's chief work, the prestimulus intervals are one to ten seconds, without foreknowledge. In Woodrow's chief work, the prestimulus intervals range from one to twenty-four seconds with foreknowledge; that is, the subject knows that the succeeding prestimulus interval will be the same as the last, till a change is announced. Woodrow's procedure gives shorter and more regular times than Breitwieser's. In the present study, the conditions of expectations are more favorable than Breitwieser's, less favorable than Woodrow's. The subject does not know what the prestimulus interval is going to be, but it is actually always close to three seconds or one second. This is more favorable than a lack of foreknowledge with a possible range up to ten seconds. It is probably also more favorable than foreknowledge of long prestimulus intervals such as 8 seconds or more. It is not so favorable as foreknowledge of prestimulus intervals of the order of those used in the present experiments. Owing to this difference in expectations, the present results are rather complementary than strictly comparable to the results of either Woodrow or Breitwieser.

7. Working with unpredicted prestimulus intervals, over a range of 1 second to 10 seconds, Breitwieser found only small differences between the various intervals. The most favored, 2 seconds, is 131 sigma, the least favored, 9 seconds, is 136 sigma. The most favored prestimulus interval for light is 4 seconds, 177 sigma, the least favored 1 second, 215 sigma, which is outstandingly long. His mean variations run 10–14 percent of their averages. A mean variation for a prestimulus interval, when large, indicates that the attention curve is especially subject to *fluctuation* at that point; when small, that at this point the attention curve is usually just about a certain height. Very distinct differences of this kind are hardly to be expected in the given conditions; it is however fairly clear that under these conditions it is the one-second interval that fluctuates the most, for both sound and light stimuli.

8. In Woodrow's work, with predicted prestimulus intervals, there is much more difference between the reaction

times for the different intervals, and they are in turn more constant. Two seconds is again most favored; one second is next. Mean variations run some 10 percent of their averages. As with Breitwieser, the most *fluctuation* is shown in the 1-second interval. The amount of fluctuation for the successive prestimulus intervals grows at first smaller and then larger again, the smallest amount of fluctuation being found at the 12-second interval. It is perhaps a freak of chance that at this point also the amount of variation shown by the three subjects differs by a mean of but two tenths of one percent. Absolutely, the fluctuation is greatest for the longest prestimulus intervals, which have much the longest reaction times.

9. The following figures represent the amount of fluctuation in the attention curve at the different prestimulus intervals. They are the average ratios of the reaction time averages to their mean variations (percent which the m.v. is of the average), and the mean variation of this average ratio for the three subjects.

Fluctuations in Degree of Readiness at Different Prestimulus Intervals (Derived from Woodrow)

	Prestimulus Interval (Seconds)							
	1	2	4	8	12	16	20	24
Fluctuation av.	11.8	10.5	8.9	9.6	8.6	10.9	9.7	10.7
M.v.	1.6	3.3	1.0	.4	.2	.9	.8	.7

10. With unpredicted prestimulus intervals ranging from 3 to 7 seconds, and from 4 to 20 seconds, Woodrow found in both cases the shortest interval to be the least favorable. This further indicates the important rôle played by the conditions of expectation in the attention adjustment. Woodrow believes that unpredicted prestimulus interval has a generally unfavorable effect on the shortest interval.

11. Turning now to the present experiments, their conditions of expectation have been described. The subject does not know what the prestimulus interval will be, but it always

approximates 1 second or 3 seconds. How the reaction times under these two intervals compare, is expressed by a ratio, the percent which the three second interval is of the one-second interval. When this figure is above 100, the one-second interval is the more favored, and vice versa.

12. The general average ratios for a group of 13 subjects (the "A-group," 216 observations each, 216 × 13 = 2,808) compare as follows with those of two subjects (11 experiments each, 216 observations, 11 × 216 = 2,376, total observations 7,560):

Percent which Reaction Times with 3-second Prestimulus Interval are of Those with 1-second Prestimulus Interval

	A-group	K	W
Av.	96	100	98
M.V.	4.5	.7	2.6

13. There is shown general tendency to favor the three-second interval, though in subject K the two are similar; the range of the A-group is from 85 to 104. This is what might be expected from previous work, since one second has appeared to lie rather ahead of the most favored interval, and three seconds, while perhaps slightly beyond it, is still in the most favored 'zone.' The conditions of expectation unfavorable for this interval, described by Woodrow, so far as they obtain here, now affect rather the one second interval. One second represents an interval during which attention has not had time to reach its maximum; at three seconds it is nearer the maximum, but has probably slightly passed it. Neither K nor W show practise effect in this function of the experiment.

14. Mention has been made of evidence that the prestimulus interval may act differently upon sound and light stimuli. Subject K alone shows this clearly, and the most direct statement of it is in the following figures:

PERCENT WHICH THE THREE-SECOND INTERVAL IS OF THE ONE SECOND INTERVAL IN SOUND AND LIGHT REACTIONS (SUBJECT K ONLY)

	Sd. 1st Half	Sd. 2d Half	Lt. 1st Half	Lt. 2d Half
Av.	104	105	95	97
M.V.	3.6	3.6	2.9	3.5

15. The apparent equality of the two intervals in K (pars. 12, 13) is seen to be the resultant of favoring the one-second interval for sound and the three-second interval for light. The other subjects do not show this, and no indication is furnished of its being a general phenomenon, or a product of the present experimental conditions. The generally better coördination of movement to sound than to light would make the result easier to interpret as a generality than as the idiosyncrasy it appears to be.

16. Differences between the two prestimulus intervals represent the height of the attention curves at those times. If the one second is favored, it indicates a rapid summoning of readiness (though not necessarily to a special height) which soon falls off. Favoring of the three-second interval indicates a slower and (perhaps) better sustained summoning of attention. The *quickness* with which attention is summoned is thus indicated by the most favored prestimulus interval; the *height* to which it is summoned by the comparative speed of the reaction at this period. Inquiry is made as to possible relation between the quickness with which attention is summoned and the height which is reached.

17. The question is if faster reactions are related to a favoring of the one-second or the three-second prestimulus intervals. The more favored interval is indicated by a ratio (par. 11); the more this ratio is below 100, the more the three-second interval is favored. This ratio, for all observations, is correlated with the central tendency for speed in all observations. For the A-group the rank difference is $\rho = -.27$ (r, 28). If sound reactions alone are considered, this figure becomes smaller, $\rho = -.22$ (r, 23) (Scott Co. Tables). Accordingly what tendency there is, is for the *quicker* summoning

of attention to go with higher summoning, as indicated in generally shortening the reaction time.

18. This does not apply to performances of the same subjects (K and W) on different days. No correlation between these factors can be asserted for the records of subject W, and with K an opposite situation appears. For all his reactions, the ρ of the two factors is .31 (r, 32), and for sound reactions only, .50 (r, 52). This individual tends to react more rapidly when he favors the three-second interval, but it is to be noted that his variations in this respect are quite slight (par. 12).

19. In par. 7 was raised the question of the *fluctuation* in the height of the attention curve at any given point after the warning signal. This is expressed in the variability of the reaction times within their unit-series. In the present experiments the criterion of such variability is the percent which the *upper* quartile is of the median (for each unit-series of 27 reactions), termed the "quartile ratio."

As these figures throw light on the dispersion of the measures in general, as well as on the particular question of fluctuation, they are quoted in some detail as follows:

Percent which Upper Quartiles are of Medians in the Various Reaction Procedures

	A-group		K		W	
	Av.	M.V.	Av.	M.V.	Av.	M.V.
Sound, 1″, first half	114	6.0	114	4.5	111	3.6
Sound, 3″, first half	114	7.9	112	3.5	111	4.2
Light, 1″, first half	113	5.8	108	2.6	108	4.3
Light, 3″, first half	110	2.7	108	1.9	106	1.8
Sound, 1″, second half	113	3.4	112	4.1	113	3.2
Sound, 3″, second half	114	4.3	111	2.6	115	3.8
Light, 1″, second half	116	7.5	108	3.5	109	2.7
Light, 3″, second half	111	3.4	108	1.8	107	2.5

20. There seems, for the light stimulus, to be very slightly more fluctuation at one second than at three seconds. The results are, however, characterized rather by uniformity. Under the present conditions, the attention curve is hardly

if at all subject to more fluctuation at one point than at another. That it is ofter *higher* at certain points than at others, has been, of course, abundantly attested.

21. This is the concluding one of three brief reports of recent simple reaction observations with normal individuals at the McLean laboratory. As already remarked, they were undertaken for comparative studies with pathological cases, with whom more extensive observations have been made. The first report of these latter is in press (4).

REFERENCES

1. Todd, J. W. Reaction to Multiple Stimuli. *Arch. of Psychol.*, 1912, 25, pp. 65.
2. Breitweiser, J. V. Attention and Movement in Reaction Time. *Arch. of Psychol.*, 1911, 18, pp. 48.
3. Woodrow, H. The Effect upon Reaction Time of Variation in the Preparatory Interval. Psychol. Monog. 1914, 17, pp. 16–65.
4. Wells, F. L., and Kelley, C. M. The Simple Reaction in Psychosis. *American Journal of Psychiatry.*

ON ATTENTION AND SIMPLE REACTION

From the Psychological Laboratory of McLean Hospital[1]

F. L. WELLS, Ph.D.
in charge

C. M. KELLEY, M.D.
Associate

GARDNER MURPHY, A.M.
Assistant

1. The variations of the simple reaction time to a constant stimulus are generally conceived as expressing fluctuations of "attention." Attention is thus used to denote a condition of the nervous system which, when it is 'high,' allows the simple reaction process to run off quickly, and when it is low, more slowly. The degree of 'attention' at any given moment of stimulus is given in the speed with which the reaction takes place.

2. Many attempts have been made to find how certain conditions of stimulation affect attention as measured in this way (1). These center about the control of an interval elapsing between a warning signal and the stimulus for reaction; here termed the prestimulus interval. After the warning signal is given, the curve of attention follows a course whose variation can be plotted for different individuals and for different expectations. If the subject knows a stimulus is coming in four seconds, his attention will be higher at that moment than if he does not know whether the stimulus is coming in one or ten seconds.

3. There seems however to be a narrow zone within which attention is relatively higher than for other prestimulus intervals. This zone of most favored intervals may be stated as from two to four seconds, with a leaning towards two. Shorter and longer intervals than these are generally less favorable, *i.e.*, give longer reaction times.

[1] As of August 1, 1920.

4. Outstanding pieces of work on this problem have been done by Breitwieser (2) and by Woodrow (3). Breitwieser gave some study to the comparison of sound and light stimuli in series where the subject did not have foreknowledge of the prestimulus interval. His general results on this point are given in the following figures derived from those quoted by him:

	Prestimulus Interval, Seconds									
	1	2	3	4	5	6	7	8	9	10
Percent which the average for the given prestimulus interval is of the average for all prestimulus intervals:										
Sound Reactions...	100.3	97.8	98.3	99.8	101.2	101.6	98.1	100.3	101.9	100.9
Do. for Light Reactions....	112	97	95	93	98	99	102	101	102	101
Percent which the Light Time is of the Sound Time (Sound-Light Ratio)	161	143	139	134	140	141	150	144	144	145

5. Breitwieser's work has a special significance for the present study through his having taken up reaction times to visual as well as auditory stimulus. Unfortunately the groups of subjects are not the same, six for the visual, eighteen for the auditory. In average, the sound-light ratios differ somewhat according to prestimulus interval. The one-second interval is the most favorable for sound (sound-light ratio 161), the four-second interval the most favorable for light (sound-light ratio 134). This result is specially interesting, as it indicates a different curve of attention for sound and light, or that attention to a sound stimulus is something different from attention to a light stimulus. In the writers' normal data, such a finding does not appear in group average, but is well established for one subject, K. His average sound-light ratio for the one-second prestimulus interval is 139, for the three-second prestimulus interval, 129. This is in accord with Breitwieser's averages for these intervals, 161 and 139 respectively. The light is favored by the longer interval. Apparently sound attention here "rises" more rapidly than light attention.

Journal of
Experimental Psychology

Vol. IV, No. 6 December, 1921

THE INFLUENCE OF VISUAL GUIDANCE IN MAZE LEARNING

BY HARVEY CARR

University of Chicago

For this experiment, a new type of stylus maze was constructed of such a character that the correct path and the cul de sacs can not be distinguished from each other by means of vision. The grooves for the stylus were milled as represented by the cross section diagram of Fig. 1. The stylus

Fig. 1. Cross section of maze representing two grooves.

consisted of a hard rubber handle at one end of which was attached by a ball and socket joint a circular metal disc of such dimensions that it moved very easily through the lower part of the groove. The pattern of the maze is represented by the diagram of Fig. 2. Its dimensions are 16 x 16½ in. The starting point is located at *A*, while position *B* is the goal or objective of each trial. The stylus is inserted in the groove at *A*, and the entrance is then blocked by a sliding metal strip. The stylus can not now be removed from the maze until position *B* is reached, where it may be easily lifted from the groove and again be inserted at *A* for the succeeding trial. So far as vision is concerned, all possible paths from *A* lead to the goal at *B*, while as a matter of fact all paths, with one exception, are blocked by invisible stops whose positions are represented by the short lines drawn across the grooves. With the aid of these lines, the course of the true path and the positions of the blind alleys are readily apparent

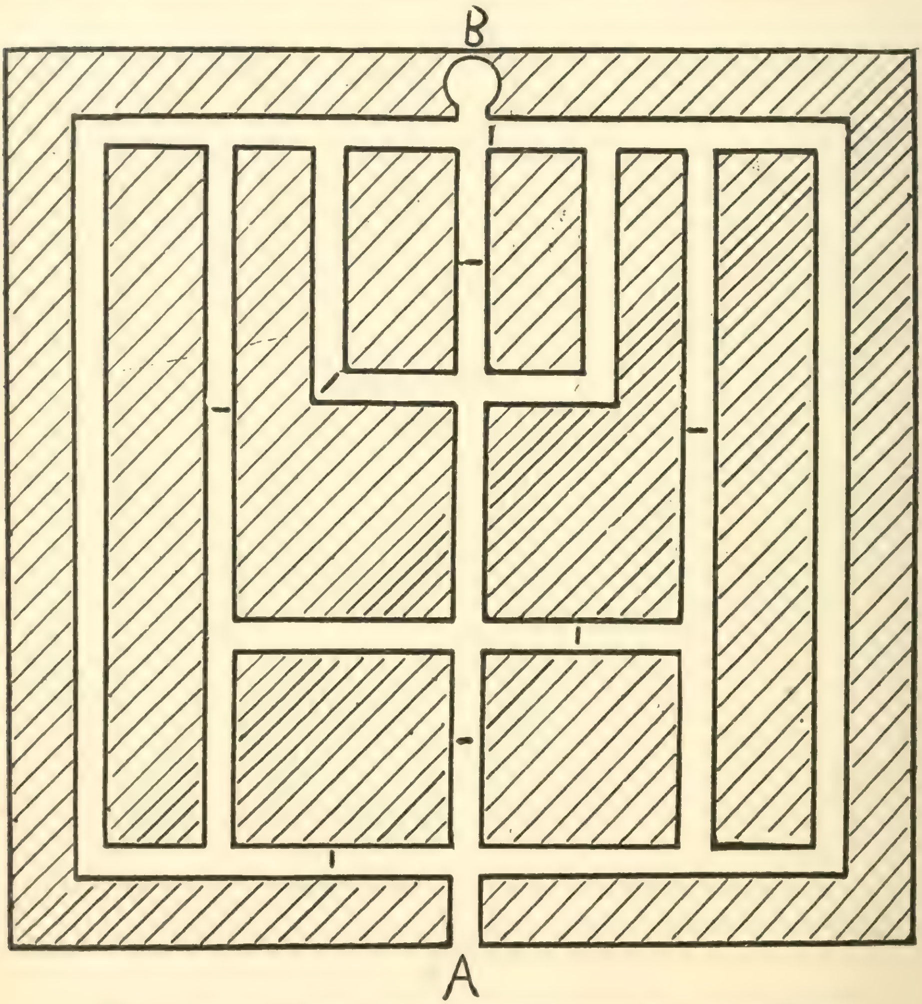


Fig. 2. Diagram of maze pattern. Lines drawn across the grooves represent the positions of the invisible stops.

to the reader. These stops project out in the lower part of the groove a sufficient distance to prevent the passage of the stylus disc, but not far enough to be visible to the subject.

The essential feature of this type of maze is the fact that the stops are invisible. As a consequence, the subject is unable to distinguish between the true path and the cul de sacs by means of vision. The maze may thus be learned in whole or part with the coöperation of visual perception, or it may be mastered in the usual manner by excluding vision by means of an intervening screen.

1. *Varying Amounts of Initial Visual Guidance.*—Six groups of subjects were employed in the first experiment. Group I. was permitted one minute for a visual inspection of the maze with instructions to obtain all the aid possible for its subsequent mastery. At the end of the period, the screen was placed in position and the maze was mastered with vision excluded. Group II. was permitted the use of vision during the first trial, and the mastery of the maze was then completed without vision. Groups III., IV., and V. were likewise allowed the use of their eyes during the initial 2, 3, and 5 trials respectively, and the learning was then completed without the aid of vision. These records are to be compared with those of Group VI. that learned the maze entirely without the coöperation of sight. This group will be termed the normal. In all cases, the maze was learned by a series of successive trials.

The comparative data are given in Table I. The first column lists the groups and the conditions of learning. In the second column is given the number of subjects employed in each group. Following are the average trial and error records for the various groups. These are total values and they include the records for both the guided and the unguided trials. The last two columns represent the influence of the visual control as measured in terms of the normal record. The positive values indicate the percentages of saving due to vision, while the negative values represent its detrimental influence. The group given five trials with vision completed the mastery of the maze with an additional 5.5 trials. As compared with the normal records, this amount of guidance was effective in producing a saving of 46.1 percent in trials and 80.9 percent in errors.

TABLE I

INFLUENCE OF INITIAL VISUAL GUIDANCE

Group	Number	Trials	Errors	Percentages of Saving	
				Tr.	Er.
I. Inspection	6	28.0	143.0	−43.6	28.8
II. 1 trial	12	20.4	86.7	− 4.6	56.9
III. 2 trials	12	19.0	95.0	2.5	52.8
IV. 3 trials	11	11.0	30.5	43.6	84.9
V. 5 trials	12	10.5	38.6	46.1	80.9
VI. Normal	11	19.5	201.2	——	——

The one-minute period of visual inspection of the maze exerted a detrimental effect upon the mastery of the problem when measured in terms of trials. It exerted a beneficial effect upon the error record, reducing the number of errors per subject by 28.8 percent in spite of the increase in the number of trials.

One and two trials with vision were approximately equal in effectiveness. They exerted no appreciable influence upon the total number of trials necessary to learn the maze, but these amounts of guidance were effective in decreasing the number of errors by over 50 percent.

Three and five trials with vision are apparently equal in effectiveness. They decreased the number of trials by at least 40 percent and the number of errors by 80 percent.

Three and five trials with vision are much more effective than either one or two trials. Obviously the effectiveness of visual guidance is not proportional to its amount. One and two trials with vision exert a much more favorable effect than does one minute of visual inspection. This amount of time devoted to visual inspection is approximately the same as that secured when vision is permitted during the first trial.

In all cases, the visual control is much more effective in preventing and eliminating errors than in reducing the number of trials.

Vision operates to prevent error not only during the trials in which its use is permitted, but also during the post-visual period of learning. The relative effectiveness of the control during the visual and post-visual stages of learning may be determined from the data of Table II. The second column gives the total error decrease during the trials in which vision was permitted. These values indicate the effectiveness of vision in preventing error in absolute terms. The third column states the average decrease per trial during the visual period. The fourth column gives the total decrease relative to the corresponding normal record in percentage terms. Vision operated to prevent 90 percent of the normal amount of error during the first three trials. The same values for the post-visual period are given in columns 5, 6, and 7 respectively.

TABLE II

ERROR DECREASE DURING VISUAL AND POST-VISUAL PERIODS

Group	Visual Period			Post-visual Period		
	Total	Ave.	Percent	Total	Ave.	Percent
II. 1 trial.......	27.3	27.3	70	87.2	4.5	54
III. 2 trials......	74.0	37.0	80	32.2	1.9	30
IV. 3 trials......	104.4	34.8	90	66.3	8.3	77
V. 5 trials......	122.2	24.4	86	40.4	6.5	68

Visual perception is extremely effective in preventing and eliminating error. When vision is permitted for one to five trials, the normal error record is decreased by 70 to 90 percent. The total number of errors prevented is not proportional to the number of trials. This result is due to the fact that the normal error curve falls very rapidly during the early trials. Stated in relative or percentage terms, the effectiveness of vision increases up to the third trial. The decreased value for five trials is due to the fact that with visual guidance practically all errors have been eliminated by this time.

There is no definite relation between the amount of visual guidance and the total saving of errors in the post-visual period. This fact may be accounted for in several ways. The post-visual periods differ considerably in length. With three and five trials of guidance, a fewer number of trials are required to learn the problem (Table I.). With these amounts of guidance, the post-visual period begins at a later stage in the mastery of the problem. When but one guided run is permitted, its post-visual effects are exerted at a time when normally many errors are being made. Measured in terms of total errors saved, any factor will be more effective in the initial stage of learning than during the final runs when the mastery of the maze is practically completed.

The influence of differences of length of the post-visual period may be eliminated by comparing the four amounts of guidance in terms of the average saving per trial. A comparison in terms of the percentage of saving will eliminate the influence of differences of position. It is apparent that the post-visual effectiveness of guidance is not proportional to its amount when measured in average or percentage values. As in the case of trials, however, three and five guided runs are considerably more effective than one or two.

Measured in relative terms, the immediate effect of guidance is invariably much greater than its subsequent effect upon the post-visual stage of learning.

Similar results were obtained from a study of the error curves whose graphs are not given. The curves for the guided groups start at a considerably lower level and fall at a somewhat faster rate than that of the standard group. With the removal of the guidance, the curves immediately rise and then descend in an orderly manner. The post-visual effect of two guided runs is confined to five trials. The curve remains below that of the normal from the third to the eighth trial, and from this point on the two are practically identical. The post-visual effect of one guided run is more prolonged. Its curve remains below the level of the normal up to the eighteenth trial. The post-visual curves for three and five guided runs are quite similar. They consistently maintain a lower level than those of the groups given one and two guided trials. In general, three and five guided runs exert the most effect upon the post-visual stage of learning, while the least effect was exhibited by the guidance given for two trials.

The immediate effect of the guidance upon the error score is easy to explain. Its effects are to be explained in part in terms of visual perception and in part in terms of a memory of these visual data. Vision permitted the subject to locate the goal and make more systematic explorations in the attempt to reach it. After traversing a certain distance in the maze, vision enabled the subject to foresee that certain paths could not lead to the goal and as a consequence certain cul de sacs were never entered. Vision enabled the subjects to identify the blind alleys more readily. Without vision, a person may enter and emerge from a short cul de sac without any knowledge of its character. Vision reduced very materially the amount of retracing over the true path, and the number of entrances into a given cul de sac during any trial. Some retracing did occur occasionally. Subjects sometimes retraced in order to explore certain paths which had been previously passed without being entered. Even with vision a subject may become lost during the first trial and retrace as a consequence. The memory factor is apparently quite effective in reducing the number of errors in the succeeding trials.

There are two ways in which vision may be effective in the post-visual period. We may assume the retention of some memory of the visual appearance of the maze pattern and its subsequent utilization as a guiding factor. The second conception assumes that the act is learned in tactual and kinæsthetic terms. Visual perception is directly effective in preventing error and thus aiding the development of a more perfect tactual motor coördination, and it is this more highly perfected act which is carried over into the post-visual period. Subjects, who have been permitted the use of vision, enter the post-visual period at a more advanced stage in the learning process than do those persons to whom vision was denied. On this assumption, the groups given guidance will be able to complete the mastery of the problem with less time and effort than the normal group.

Both conceptions are plausible. They are not mutually exclusive alternatives and both may be true. No facts were obtained which prove or disprove the second assumption. Several lines of evidence support the validity of the first hypothesis.

Visual inspection of the maze enabled the subjects to localize the goal and obtain some notion of the size and character of its pattern. This preliminary visual experience did not, however, permit any differentiation between the true and false paths. The effects of this inspection were retained and it exerted some influence upon the process of learning. Its influence upon trials was detrimental, but the total error record was decreased by 28.8 percent and the average number of errors per trial by 50 percent. The facts prove rather conclusively that a memory knowledge of the visual features of the maze can be utilized as a guiding factor.

The fact that the effectiveness of the visual control is not proportional to its amount can not be explained wholly on the basis of the tactual motor hypothesis. According to this conception, visual inspection should exert no effect, five guided runs should be more effective than three, and two trials more effective than one. The fact can be explained in terms of a memory of the visual data. Both detrimental and beneficial effects will result from an inadequate and partially incorrect notion of the appearance of the maze pattern. The beneficial effects will increase as the guiding concept becomes more definite and precise, and the maximal amount of benefit will be obtained from that amount of guidance necessary to develop a correct conception of the maze pattern. As a matter of fact, the maze was visually mastered at the end of the third trial by the majority of the subjects. Of the first two groups, but one of the twenty-four subjects obtained a visual mastery of the maze within two trials. When three guided trials were given, seven subjects were able to identify the correct path by sight. With five guided runs, seven subjects mastered the problem in terms of vision by the end of the third trial, and three others by the end of the fifth trial. The average number of trials necessary to learn this maze in visual terms is probably somewhere between three and four. Since a correct notion of the maze pattern is attained in about three trials, any additional visual experience will be without effect.

The effectiveness of the memory factor was indicated by the remarks of the subjects. Its efficacy apparently varied with the individual. Some asserted that they retained an adequate memory of the appearance of the maze pattern and that the few errors made were due to carelessness. This type of subject traversed the maze slowly and carefully and made the best records. Some subjects exhibited a certain degree of uncertainty and indecision at the beginning of the first trial without vision. These people asserted that this uncertainty was due to the lack of any adequate memory of the visual path. Others did fairly well on the first trial but exhibited considerable hesitation at the beginning of the second trial. These subjects ascribed this hesitation to the fact that the appearance of the maze had been forgotten in the meantime. In other cases the subjects proceeded confidently until difficulties were encountered, and these mistakes were usually attributed to a wrong visual conception. A few subjects reported that this memory control was detrimental inasmuch as they discovered after several trials that they could do better by relying exclusively upon kinæsthetic clues for guidance.

After the maze was mastered without the use of vision, all groups were required to run it for several trials when vision was not excluded. All subjects in the groups given three and five trials of visual guidance made no errors in this test. Evidently these subjects had developed a correct notion of the appearance of the maze pathway, or their movements were guided wholly in kinæsthetic terms. In the two trial group, two subjects made one error each which they ascribed to an incorrect idea of the nature of the path. In the one trial group, one subject had developed a very erroneous concept of the maze as a visual object and as a consequence made twenty-two errors in his first attempt to traverse the maze with the use of vision. One error was made in the second attempt, but the third was successful. Vision was now excluded and four errors resulted. Evidently the correction of the visual control had disturbed the kinæsthetic coördination. Three of the six subjects who had been given one minute for a preliminary inspection of the maze pattern made slight mistakes when attempting to traverse it with perceptual guidance. The test was also given to the group that learned the maze entirely without the aid of vision. These subjects invariably reported that the maze pattern looked very much different from what they had anticipated. Two individuals made slight mistakes which were apparently due to the fact that the introduction of vision increased speed. Two other subjects made errors because they were guided in part by vision.

Table III

Influence of Position of Guidance

Group	Guidance	Total Saved				Percentage Saved			
		Tr.	Err.	V.	P.V.	Tr.	Err.	V.	P.V.
III. 12........	1–2	.5	106.2	74.0	32.2	2.5	52.8	80	29
VII. 10........	5–6	3.3	53.4	16.7	36.7	17.0	21.0	75	67
VIII. 12........	9–10	− .5	17.2	.2	17.0	− 2.5	8.5	3	46

2. *Position of Visual Guidance.*—Two trials of visual guidance were introduced at different stages in the learning process. Group III. was permitted to utilize vision during the first two trials and was then required to master the maze without this aid. Group VII. ran the first four trials without vision. Visual guidance was then introduced during the fifth and sixth trials, and after this the mastery of the maze was completed with vision excluded. With group VIII., the visual guidance was introduced during the ninth and tenth trials. The comparative data are given in Table III. The results are stated in terms of the effects of the guidance relative to the records of group VI. that mastered the maze entirely without the aid of vision. Group VII. consisted of 10 subjects. Guidance was given during the fifth and sixth

trials. This group learned the problem in 3.3 fewer trials than did the normal, and made 53.4 fewer errors in so doing. Of these errors, 16.7 were saved during the period of guidance and 36.7 during the post-visual period of mastery. The values in the last four columns represent the same data stated in terms of percentages. For example, the saving of 16.7 errors which occurred during the two guided runs is 75 percent of the error record of the normal group for the corresponding trials.

Visual guidance introduced in the initial or final stages of learning exerts no appreciable effect upon the number of trials required to master the maze. Two guided runs in these positions are no more effective than an equal number of unguided trials. When two trials of visual guidance are inserted in the intermediate position, they exert a greater effect than a similar number of unguided trials.

Visual guidance for all positions is much more effective upon errors than trials. The earlier the introduction of vision, the greater is its effect upon the total error score.

Guidance in all cases is effective upon errors during both the visual and the post-visual stages of learning.

Considering the *immediate* effect of visual guidance upon errors, it is evident that the earlier its introduction, the greater is its effect when measured either in absolute or percentage terms. For the initial position, vision reduced the error score for all subjects. In the intermediate position, vision was effective only with those subjects who had made but little progress in error elimination. The introduction of vision in the final position was slightly effective with a few subjects, not effective with the majority of the individuals, and detrimental with two subjects. This distractive effect was apparently due to the fact that these individuals had developed an erroneous notion of the visual appearance of the true pathway, and consequently became confused and lost when vision was introduced.

The effect of vision during the post-visual period is greatest for the intermediate position when measured in terms of the total errors saved, the average saving per trial, and the percentage of saving. The initial position is more effective than the final position in respect to the total and average number of errors saved. Measured in terms of the percentage of saving, guidance in the final position is the more effective. The saving resulting from visual guidance in the initial position is confined to the first five trials of the post-visual period. Guidance in the other positions manifests an effect throughout the post-visual period. The post-visual efficacy of guidance apparently varied with the individual. Guidance in the initial position was quite effective with some subjects and detrimental with others. When introduced in the intermediate position, guidance was effective in all cases. The efficacy of the control for the final position was confined to a few individuals.

No generalization can be made as to the most advantageous position in which to introduce a visual perceptual control. It is possible that more advantageous results might have been secured by the introduction of vision during the third and fourth trials. Moreover, the most effective position may vary with the amount of guidance given. The character of the problem may also be a conditioning factor.

3. *Influence of Indirect Visual Guidance.*—A final experiment tested the influence of indirect visual guidance by the use of map diagrams representing various features of the maze pattern. These diagrams were exact duplicates of the maze as to size. They were placed on the table at the left of the screened maze. The nature of these maps was explained to the subject and he was instructed to utilize the diagram in guiding his movements through the hidden maze.

Group IX. mastered the maze with the use of map *A*. This diagram represented all the possible pathways but it did not enable the subject to distinguish between the true path and the cul de sacs. It represented the maze pattern as it appeared to vision. Group X. learned the maze with the aid of map *B* which is represented in Fig. 2. This diagram was like that of map *A* with the single exception that the positions of the various stops were indicated by lines drawn across the pathways. These lines enabled the subjects to distinguish between the cul de sacs and the true path. Group XI. utilized map *C* which represented the course of the true path. The subject was informed that the maze contained numerous cul de sacs but that these were not represented in the diagram. The comparative data of these groups are given in Table IV. in addition to the records of the normal group that mastered the maze without the aid of vision.

Table IV

Influence of Indirect Visual Guidance

Group	Number	Total Trials	Initial Errors	Total Errors	Percentage of Saving	
					Trials	Errors
VI. Normal......	12	19.8	54.1	201.2		
IX. Map *A*.......	11	11.8	47.5	105.4	40	48
X. Map *B*.......	11	7.5	3.7	14.7	62	93
XI. Map *C*.......	11	14.5	5.2	31.2	27	84

All three modes of indirect visual guidance were extremely effective in reducing both the number of trials and the amount of error requisite to the mastery of the problem. In all cases the controlling factor was more effective upon errors than trials.

The effectiveness of the guidance upon trials and errors was proportional to the amount of knowledge obtained from the diagram. Diagram *B* was the most effective, and obviously it gives the most information concerning the maze. Likewise, more information can be derived from diagram *C* than from map *A*.

There is no evidence that map *A* was effective in reducing the number of errors during the first trial. The difference between this value and that of the normal group may be accounted for in terms of chance. Moreover, this form of guidance was about equally effective upon both trials and errors. The error curve for this group begins approximately at the same height as that of the normal group, but it drops very much more rapidly. This form of guidance exerts its main effect after the subject acquires some notion of the maze pattern during the initial trial.

Diagrams *B* and *C* were very effective during the first trial. The initial error records of these two groups were 3.7 and 5.2 respectively. The two curves maintain a similar relation to each other throughout the mastery of the problem. Both forms of guidance were very much more effective upon errors than trials.

The initial influence of these modes of indirect guidance may be compared with the direct effect of vision. The average error record for the initial trial of the four groups that were permitted the use of vision was 11.7. These conditions were like those obtaining for the group that utilized map *A*, with the single exception that the guidance was introduced directly in the one case and indirectly in the other. Direct guidance is the more effective not only during the initial run but throughout the learning process. The indirect guidance exerted by diagrams *B* and *C*, however, is more effective during the first trial than is the direct influence of visual perception.

With map *A*, the subjects were required at the end of each trial to identify the course of the true path in terms of the diagram. At no time were the subjects informed as to the correctness of these judgments. The learning process with this mode of guidance may thus be divided into two stages,—the initial period during which the pathway is being visually identified, and the final stage of completing the mastery of the problem. The records for the two periods of learning are given in Table V.

Table V

Two Stages of Learning with the Use of Map *A*

	Trials	Initial Error	Total Error	Average Error
Initial period	3.8	47.5	92.0	24.2
Final period	8.0	2.0	13.4	1.7

The first identification was incorrect with each subject. Three of the eleven subjects were successful after the second run. The average number of trials required by the group to master the maze in visual terms was 3.8, and the average error record per trial for this period was 24.2.

A visual mastery of the maze did not enable the subjects to traverse it without error. Eight additional trials were required. Individuals differ in their ability to guide their movements by a visual diagram. Two subjects were able to run the maze correctly as soon as they succeeded in identifying the pathway on the map. A third subject made 2 errors in his first attempt but was successful thereafter. A majority of the individuals required a considerable number of trials to complete the mastery of the problem. Few errors, however, were made in the final period. Of the 88 trials occurring during this period, an error score greater than 3 occurred in but three runs. The errors were slight but difficult to eliminate.

Similar results obtained with the use of map *B* which enabled the subjects to identify the course of the true path before attempting to run it. One subject was able to traverse it successfully on his first attempt, while another subject mastered the problem in one trial. Other individuals manifested considerable difficulty in guiding their movements by means of the diagram, but in all cases very few errors were made in any trial.

The subject's confidence in his identification was in general proportional to its correctness. The first confident judgment was invariably correct, and with the exception of one subject a correct identification was never subsequently revised. This individual correctly and confidently identified the true path at the end of the third trial. This judgment persisted for the fourth and fifth runs, and a wrong path was then chosen at the end of the sixth trial. The subject reverted to his previous correct choice on the next trial. This judgment was again revised after the ninth run and this wrong conception persisted until the problem was completely learned. We have here the anomalous case of the development of the proper kinæsthetic motor coördination under the supervision of a wrong visual conception. The record of this subject, however, was not unusual in other respects. His trial and error scores were somewhat above the average for the group. Poorer scores were made by other individuals, and the deviation from the group average may well be accounted for in terms of chance. His error of identification was slight, however. When required to run the maze with the aid of direct vision, this individual made but one error in his first attempt.

During the initial period, the subject acquires two factors which are potentially significant for the subsequent mastery of the problem. The correct identification of the true pathway on the diagram is one. The second factor is the kinæsthetic motor coördination which has been developed during these trials. The relative value of the two factors may be determined from a comparison of the records for the final period with those obtained by the group that used map *B* as a guide. Group IX. learned to identify the true path on map *A* after 3.8 trials. Group X. was furnished this means of identification by map *B*. Group X. thus began the problem at a point comparable with that attained by group IX. upon entering upon the final period of mastery. Any difference in the two records must be due to the influence of the kinæsthetic-motor tendencies which were developed during the initial period. The comparative data are given in Table VI.

Table VI

Influence of Visual Identification of the True Path

	Trials	Initial Error	Total Error	Average Error
Map *A*, final period	8.0	2.0	13.0	1.7
Map *B*	7.5	3.7	14.7	1.9

The two sets of records are practically identical in every respect. We are thus forced to conclude that the motor activity involved in running the maze during the initial period was primarily valuable on account of its visual consequences. Presumably certain kinæsthetic motor tendencies were also acquired, but apparently these tendencies must have exerted both detrimental and beneficial effects upon the subsequent process of learning.

After the problem was learned, certain subjects were required to run the maze without the use of the diagram. Five individuals of group IX. were subjected to the test. Two ran the maze perfectly for several successive trials, one made a slight error in his first attempt, while two experienced some difficulty for several trials. Ten subjects of group X. were tested. Six ran the maze perfectly. Three made a few errors for several trials, while one individual experienced considerable difficulty in mastering this new phase of the problem. He required 12 trials and made 17 errors in learning to adjust to the new situation, as compared with 13 trials and 30 errors in learning to run the maze with the aid of the diagram. Evidently individuals differ very materially in their reliance upon the perceptual control after the problem is learned. The ability to run the maze without the aid of the diagram was apparently not correlated with the number of trials required to learn the problem. In this connection it was noticed that several subjects dispensed with the perceptual use of the diagram when the problem was about half mastered. In response to inquiries they replied that they preferred to rely upon a memory of the diagram as a control. Others reported that it made no difference whether they perceived or remembered the diagram in the later stages of learning. On the other hand, some individuals felt that the perceptual means of control was essential to a good record.

When the diagram was removed after the problem was mastered, the maze was traversed in terms of kinæsthesis or by means of a memory of the diagram. We have no decisive data upon which to base an estimate as to the relative importance of the two factors. Individuals differed in their subjective estimates of the importance of the memory factor, but such reports can not be taken seriously.

4. *Influence of a Visual Control derived from Kinæsthetic Data.*—Although any visual perceptual contact with the maze was denied the normal group, it must not be supposed that all modes of visual guidance were excluded. Under these conditions, subjects do gradually construct from their unseen movements some sort of an idea of the visual appearance of the true path. It is reasonable to suppose that this conceptual schema as it develops operates in turn to influence the succeeding attempts at traversing the maze, for our previous experiments have indicated the effectiveness of such concepts when derived from visual perceptual data.

The question as to the influence of such a control can be answered by a comparative study of the records of blind and normal subjects. I am indebted to Miss Helen Koch for the data given in Table VII. which were obtained with the use of a different maze from that employed in our experiments. The normal group consisted of ten college students. The table gives their average scores in trials, total errors and total time, in addition to the poorest record obtained for each of these three measures. The three subjects with defective vision were also college students. Subject *A* was able to see shadow differences in the visual world. Subjects *B* and *C* have been totally blind from birth. Subject *B* was a Phi Beta Kappa student, and subject *C* was able to enter the Law School after graduation from college. The records were obtained during their undergraduate work.

TABLE VII

COMPARISON OF NORMAL AND BLIND SUBJECTS

	Trials	Total Errors	Total Seconds
Normal average	44.3 ± 21.8	325 ± 197	$1{,}466 \pm 766$
Poorest normal score	97	740	4,112
Subject *A*	46	566	1,731
Subject *B*	92	2,705	9,595
Subject *C*	83	3,671	7,291

There is no conclusive evidence that the ability of subject *A* was limited by his visual defect. His record in every case was less than the poorest normal score. While his values were greater than the corresponding normal averages, yet in only one case do they exceed the average range of normal variability.

The totally blind subjects were evidently at some disadvantage in mastering such an act of skill. Their scores for trials exceed the average range of normal variability, but lie within the maximal range. Both the error and time values, however, lie far beyond the maximal range of normal variability. The error records are four to five times larger than the poorest normal score, while the time values are about twice that of the largest normal record. The visual deficiency of these subjects has certainly affected their time and error records, and possibly it has influenced their trial scores to a slight extent.

More comparative data are necessary for any confident generalizations. Moreover, it is somewhat of an assumption to maintain that the totally blind subjects were at a disadvantage primarily for the lack of a capacity for visualization. Granted the assumption, however, it is to be noted that the results confirm our former conclusions that the indirect mode of visual control is much more effective in reducing errors than trials.

Our experiments have been concerned with a general problem of considerable theoretical and practical importance, viz., the significance of the visual processes in the acquisition and performance of various acts of skill. As a general rule, both the direct and indirect participation of vision in these activities has been largely ignored. It has been the usual custom to describe such activities wholly in tactual and kinæsthetic terms. Our results at least indicate the probability that the visual modes of control are much more effective in the acquisition and performance of acts of skill than hitherto suspected.

SMALLER VS. LARGER UNITS IN LEARNING THE MAZE

BY J. W. BARTON

University of Idaho

Investigation has raised a doubt in the minds of some[1] concerning the validity of the statement that to 'learn by wholes rather than by parts' [2] is the more efficient method. Whenever such experimental results have been presented the investigator has attempted an explanation of why learning by smaller units should be found more efficient as a learning means.

Continued investigation with respect to this matter seems to give stronger and stronger evidence that for maze learning (motor learning) [3] the smaller unit methods are not a little, but are very much, better as learning means than is the case for very large ones.[4] Similar results are being found for other forms of learning.

The data for this study were obtained during the second semester of the school year 1920–21. The subjects used were members of one of the psychology classes made up of six males and twenty-eight females. These subjects were assigned to the various method-groups by chance selection. The 'whole method' group was made up of 2 males and 10 females, the 'part continuous method' group of 2 males and

[1] Peckstein, L. A., 'Whole Versus Part Method in Learning Nonsensical Syllables,' *Jr. Ed. Psychology*, 1918, p. 387.

[2] Starch, Daniel, *Educational Psychology*, 1921, p. 185.

[3] Possibly all learning is motor and that the different types indicated by most writers are only evidences of the specific nature of the nerves involved in the respective aspects of adjustment.

[4] I appreciate the uncertainty of meaning attached to the words 'whole' and 'part' in discussions on learning. I suppose no one doubts that a unit of practice might be made too large for adequate learning for any individual in any field of study. No one can mean such a 'whole' as this in such discussion, but just what is meant has been defined for each particular case as it has come up.

9 females, and the 'part method' group of 2 males and 9 females.

For this study the meaning assigned for each of the three methods used is as follows: By the 'whole method' is meant that the learners always started at the entrance of the maze and continued in their random efforts until they reached the end of the complete maze. This was repeated in each case until the maze was learned. By the 'part continuous method' is meant that such learners began always at the entrance of the maze but first finished at the end of the first quarter. After this quarter was learned the second quarter was added to the first and these two together served as the second unit. When these two were learned the third quarter was added to the first two. When this three-quarter-unit was learned the last quarter was added. Then the whole maze was taken as a unit and completed. By the 'part method' the learners learned each quarter separately in regular order. After each one had been so learned they were combined into a single unit and relearned as a whole.

Each individual of the various groups was required to come each day at a designated time, for a period of fifteen minutes, until the learning had been completed. This was not always complied with, but the irregularities were pretty well distributed throughout. The room in which the work was done was the same for all subjects, and the position of the subjects and the experimenter in the room was rather constant. On a few occasions a subject was required to use a different room, but as far as the groups were concerned there was no difference in treatment in any matter except as the nature of the experiment required. Two of the subjects in the 'part method' group dropped out of school before finishing the experiment.

The intelligence of the various subjects, as measured by the Otis Group Scale, and the college grades for the school year 1920–21 are shown for each group by Tables I., II. and III.:

Tables Showing Results of Otis Tests and the Average School Marks for the Year 1920–21

Table I

Whole Method Learners

Subjects	School Marks	Score in Each of Ten Tests										Totals
		1	2	3	4	5	6	7	8	9	10	
I.....	5.162	10	21	22	17	16	14	18	11	22	25	176
II.....	5.470	14	25	21	18	16	12	19	16	24	21	186
III.....	3.527	12	15	4	11	5	6	8	11	17	12	101
IV.....	5.805	16	22	21	20	16	13	18	5	23	21	175
V.....	4.416	13	17	12	14	15	9	15	10	18	22	145
VI.....	5.214	11	16	10	11	15	10	12	11	23	19	138
VII.....	5.243	11	20	19	18	16	13	13	10	18	21	159
VIII.....	5.405	12	18	18	18	12	10	13	3	20	26	150
IX.....	5.089	9	18	14	16	12	10	13	10	18	21	141
X.....	5.176	12	16	6	19	8	7	7	8	24	19	126
XI.....	5.400	14	22	18	9	16	11	15	15	12	24	157
XII.....	5.342	10	18	17	12	12	13	15	11	23	25	156

Table II

Part Continuous Learners

Subjects	School Marks	Score in Each of the Ten Tests										Totals
		1	2	3	4	5	6	7	8	9	10	
I.....	5.314	12	20	17	17	15	14	18	11	23	24	171
II.....	3.875	12	19	18	12	14	10	18	15	21	24	163
III.....	4.393	14	17	21	18	18	8	9	16	24	17	162
IV.....	3.880	14	20	24	19	18	11	19	16	21	25	187
V.....	5.242	14	14	13	17	7	9	10	8	15	22	129
VI.....	5.171	12	15	14	19	17	10	14	13	18	22	154
VII.....	5.114	13	20	18	17	16	12	16	18	22	20	162
VIII.....	4.800	13	17	16	15	15	11	11	10	24	21	153
IX.....	4.566	..	..	..	..	..	..	..	..	..	..	...
X.....	3.111	9	15	22	15	10	7	8	6	23	20	135
XI.....	3.500	..	..	..	..	..	..	..	..	..	..	...

The materials used in this experiment were first a maze made up of a board 20″ x 15″ x 1″ on which were nailed cleats of board one inch wide and three-eighths inch thick in such way as to leave a groove about three-eighths inch wide. There were 31 cul de sacs ranging in length from one to four and one fourth inches and a true path of exit making a com-

Table III

Part Learners

Subjects	School Marks	Score in Each of the Tests										Totals
		1	2	3	4	5	6	7	8	9	10	
I.....	5.200	11	17	16	19	15	10	16	13	22	18	157
II.....	5.218	11	19	15	19	16	13	17	7	22	16	155
III.....	5.192	13	20	19	11	15	12	13	15	15	24	157
IV.....	3.807	12	18	23	16	10	9	14	11	17	19	141
V.....	5.550	13	16	20	17	13	12	14	11	25	24	165
VI.....	5.542	13	16	16	18	15	12	16	9	19	19	153
VII.....	4.617	10	17	7	8	14	8	9	4	22	20	129
VIII.....	5.459	9	16	16	17	11	4	15	11	21	23	143
IX.....	4.334	13	21	16	18	13	11	9	10	16	16	143

plete length of 107 inches. Through this groove of cul de sacs and true path the subject was required to trace from the entrance to the end of the groove, by means of a penholder, turned top end down, while he sat blindfolded on the opposite side of a table from the experimenter.

The maze was numbered every inch or so for purposes of keeping a record of just what part of it was of greatest difficulty. The experimenter kept a complete record of the number of errors, the place of error on the maze, the number of trials, and the time. A stop watch was used in recording the time.

A trial is a tracing from the start to the end of a designated unit. This means that the trials were longer for the larger-unit learners. The 'whole' learners had the possible maximum length of trial each time, the 'part continuous' learners had a trial the length of which became greater as the quarters were added, and the 'part' learners had the shortest trials, for a unit in this case consisted of only one fourth of the maze at a time until the fourths were combined. At this point the latter group had a trial equal in length to the 'whole' learners. The whole-maze unit, or any other unit, was considered learned when the subject succeeded in going through it three times in succession without error.

In no case was a learner permitted to see the maze until after it had been completely learned. Many of them knew nothing about the location of the end of the unit to be learned, how large the maze was, or whether they were to walk around in it or not. Some of the last ones to work on the experiment had been talked to by the ones finishing earlier, but the three groups shared about equally in this.

The instructions that were given to the subject, in each case, on beginning the experiment for the 'whole' learners were as follows: "On the reverse side of this board narrow strips of wood are nailed so as to leave a space between them. This space is made so that it is a continuous groove with many blind alleys branching from it. Your problem is to trace through this continuous groove by means of a pen handle. There is only one possible way out. A mistake is made when you go into a blind alley and then have to retrace to get out, or when you retrace a certain distance in the true path toward the beginning. A blind alley never has more than three turns in it. You will continue at the work until you can trace through the maze three consecutive times without error."

"You will continue this work fifteen minutes each day until you have completed it. Have you any questions?"

The instructions for the 'part continuous' and for the 'part' learners were the same as for the 'whole' learners, except that there were changes made to let them know how the ends of the smaller units were to be indicated. In all cases of tracing through a unit the hand of the subject was guided back to the start by the experimenter for the beginning of a new trial. Whenever a subject would retrace to the start the experimenter so signified by saying, 'start.'

Below are tables which show the number of trials, the number of errors and the time in seconds for each subject of the respective groups as well as the average for each item of measurement:

The tables for the 'part' learners and the one for the 'part continuous' learners show figures in the column of trials that are inclosed in parentheses. These are included because they

Tables Showing Number of Trials, Number of Errors and Time in Seconds

Table IV

Part Continuous Learners

Subjects	Trials	Errors	Time
I	109	364	5,520
II	73	227	4,645
III	194	1,491	6,825
IV	130	577	5,315
V	72	389	3,305
VI	83	378	3,435
VII	90	408	3,445
VIII	86	1,228	4,770
IX	84	397	5,185
X	168	568	6,015
XI	73	548	4,380
Average	105.6(52.8)	597.7	4,803.6

Table V

Whole Learners

Subjects	Trials	Errors	Time
I	30	609	3,770
II	103	1,726	6,535
III	37	992	3,863
IV	46	1,461	5,565
V	39	673	4,825
VI	35	687	3,360
VII	49	887	4,815
VIII	53	712	5,360
IX	98	1,666	5,110
X	32	1,066	4,055
XI	79	1,753	5,685
XII	39	1,189	6,670
Average	53.3	1,118.4	4,976.5

indicate more correctly the true averages for trials, if they are to be directly compared with such results for the 'whole' learners. A trial for the 'part' learners was only one fourth as long as was the case for the 'whole' learners for most of the learning practice. The trials for the 'part continuous' learners were very much shorter in the same way, except that there were more attempts at the whole maze in this method of practice than was the case for the 'part' learners.

Table VI

Part Learners

Subjects	Trials	Errors	Time
I.	97	399	3,405
II.	131	373	5,060
III.	63	692	4,390
IV.	39	349	2,810
V.	81	531	3,735
VI.	72	323	3,700
VII.	85	436	3,290
VIII.	97	358	2,840
IX.	132	836	3,715
X.	...	...	....
XI.	...	...	....
Average	88.6(22.15)	477.2	3,660.5

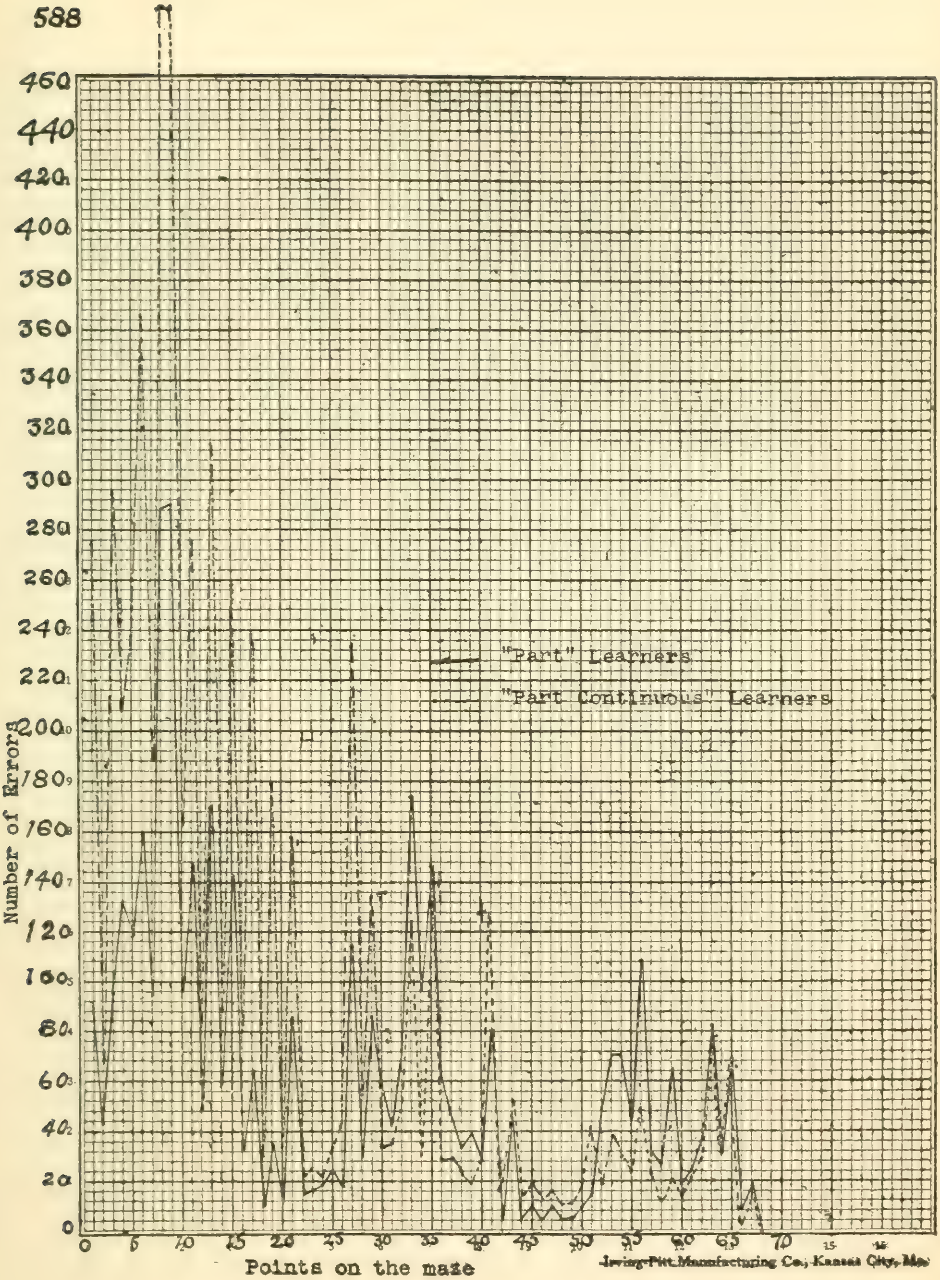
Other aspects of learning are presented by Tables VII. and VIII. and the curves that follow. These show the number of errors for each of the quarters of the maze used in this experiment. These data indicate what the case is here for such matters as 'transfer,' 'backward elimination,' and repetition as factors conditioning learning. Such material is of use in explaining why a given method proves nore effective as a learning means.

Average Number of Errors by Quarters for Each Method

Table VII

Method	First Quarter	Second Quarter	Third Quarter	Fourth Quarter
Whole	352.6	167.4	87.4	71.5
Part Continuous	386.9	112.8	54.8	43.2
Part	231.4	109.4	59.8	82.8
Totals	972.9	389.6	202.0	197.5

If the data presented in Table VII. is presented in the form of the per cent. that each other quarter is of the fourth



"Part" Learners
"Part Continuous" Learners
Number of Errors
460
440
420
400
380
360
340
320
300
280
260
240
220
200
180
160
140
120
100
80
60
40
20
0
0
5
10
15
20
25
30
35
40
45
50
55
60
65
70
Points on the maze
Irving-Pitt Manufacturing Co., Kansas City, Mo.

quarter, in the number of errors made, we get the figures given in Table VIII. below:

TABLE VIII

Method	First Quarter	Second Quarter	Third Quarter	Fourth Quarter
Whole	493.2%	233.8%	123.4%	100%
Part Continuous	896.0%	261.0%	126.9%	100%
Part	279.2%	123.8%	72.2%	100%

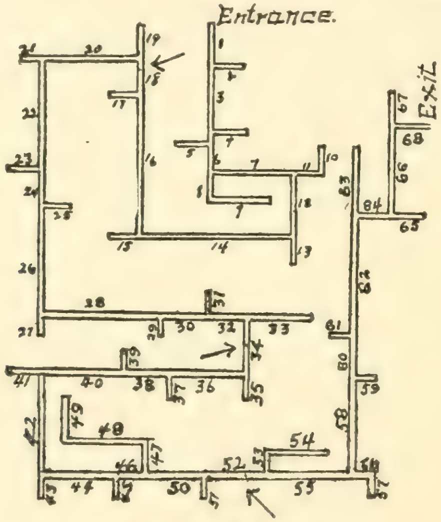


DIAGRAM OF MAZE
The arrow indicates the ends of the quarters.
A small wooden plug was inserted at these points.

Tables IV., V. and VI. show very plainly that for errors the '*part method' is very superior as a learning means*, for this exercise, when compared with the 'whole method' results. There is a gain of 57.3 per cent. in this aspect of learning when using the smaller units. The 'part continuous method' shows a gain of 46.5 per cent. over the 'whole method' as a learning method in the number of errors made.

In the matter of time, the gain over the 'whole method' is not so pronounced—it being 26.4 per cent. for the 'part method' and 3.4 per cent. for the 'part continuous.' When what was said above concerning the number of trials is considered, it will be seen that the gain for either of the 'part methods' over the 'whole method' is very considerable in all three particulars.

These differences in performance in this experiment can not be explained on the basis of difference in mentality for the three groups, for the 'whole' learners have higher school marks than either of the other two groups. In the mental tests the 'whole' learners came second in mentality as measured by this test. The 'part' learners, as a group, was the lowest of all. The differences for the Otis Tests were so slight, however, that they can not be considered as significant. We must look elsewhere for an explanation of the superiority shown by the 'part' learners in this study.

There is evidence of very great 'transfer' if Tables VII. and VIII. are observed. Here we find a gradual improvement in learning for each succeeding quarter in all three methods except for the case of the fourth quarter in the 'part method' figures. In this single case the third quarter was less difficult by 27.8 per cent. than the fourth quarter.

This evidence of transfer should be expected, for the parts of the maze are very much alike. This similarity in parts *involves much the same nerve processes* in the impulse phenomena required in the responses had in successfully reaching the end of the unit of learning used. By looking at the diagram of the maze it will be seen that the last quarter is less like the others than any other one.

What has been said about the 'positional establishments' [1] when the units are being connected, as an explanation of the waste in 'part' learning can have no point in this experiment; because the 'part method' is, by far, the best of the three methods here used. There is no waste to be explained and the 'negative transfer' [2] is not found except in going from the third to the last quarter. The loss here could likely be explained on the basis of the maze difference indicated above. Such a difference as this necessarily involves a greater dif-

[1] Peckstein, L. A., 'Alleged Elements of Waste in Learning a Motor Problem by the "Part" Method,' *Jr. Ed. Psychology*, 1917, p. 309.

[2] *Ibid.*, p. 306.

ference in *nerve functionings* than that found for the other three quarters.

Either of the 'part methods' shows advantages over the 'whole method.'[3] This was the case in this experiment even though the connections of the smaller units were not attempted until after each of the four quarters had been learned as isolated processes. This advantage is likely explained, partially at least, on the transfer basis.

Transfer seems to be best taken care of when *there is identity of neural factors involved* in the two or more situations of response; provided that these neural factors are most *ready to respond* at the different times and in the different situations of stimulation. This readiness seems to be best provided for *by the needs of the being* for organization required in response. The organization is determined by the nature and the nurture of the subject used. What is presented at any point in the accumulation, to be effectual, *must involve neural elements already utilized in this organized accumulation;* for it appears that each being is a unity at any time, however poorly directed some might seem. Learning then likely means organization in neural bond relations for response plus the tendency to act on occasion of stimulation. The stimulation situation must be of a size that will best bring into function all the matters of readiness as well as those of identity in neural units used if the transfer effects are to reach the maximum for adequate learning.

It is not shown that the parts used, in the most adequate learning of this maze, were just the proper size to obtain the best results. It is a 'lucky strike' if they are of the optimum length. The size of unit to use, for the best results in transfer, is that which involves neural factors most widely used in units to be learned subsequently; provided that it is not too large to get the maximum helps from the readiness for response factors. To make it too large is to overwhelm the

[3] Peckstein, L. A., 'Whole Versus Part Method in Learning Nonsensical Syllables,' *Jr. Ed. Psychology*, 1918, p. 386.

learner and thus prevent the adoption of the new impulse phenomena into the ever-accumulating organization.

It might be found that there is an *optimum size of unit for each kind of material to be learned*, and again it might be the case that the size of the unit should differ for each individual learner. Subject No. 6 used only 3,360 seconds in learning the whole maze by the use of the 'whole method.' This is better than any other learner regardless of the method used, except for cases No. 4, No. 8, No. 7 of the 'part learning' group. Other such individual differences within each group are shown, as well as much overlapping from group to group. Such evidence seems to imply that a general rule with respect to the size of unit to use in verbatim learning is going to be hard to define.

Notwithstanding these apparent discrepancies there seems to be ample evidence to justify the doubt that has been raised in the minds of some concerning the validity of the statement that to 'learn by wholes rather than by parts is more efficient.' [1]

[1] We have gathered data on nonsense learning and on the memorizing of prose and poetry which is being put in form for publication. While it is a little early to determine positively as to the superiority of the methods used, the indications seem to avor some form or other of the 'part method' as being the more efficient.

IMPROVED FORMS OF STEADINESS TESTER AND TAPPING PLATE

BY KNIGHT DUNLAP

The Johns Hopkins University

I. Adjustable Steadiness Tester

The form of steadiness tester illustrated in Fig. 1 has been in use in several departments of the Johns Hopkins University, and in other institutions, for several years, and has proved its value. The holes of graded sizes are arranged in a circle in a brass disc, which is placed in front of a second disc of the same size, having a single hole slightly larger than the largest hole in the first disc. The first disc turns on the short brass rod to which the second disc is fastened, so that any of the test holes can be brought over the single hole in the second disc. Two brass springs fastened to the second disc, and rubbing on the first disc, furnish friction sufficient to hold the first disc in any position, and also provide good electric contact between the two discs. By means of a thumbscrew on the rod, the pressure of the springs can be adjusted.

The rod carrying the two discs is held by a combination clamp on a vertical rod, so that the angle between the disc-surface and the line of sight can be varied; the position of the single hole in the second disc can be varied; and the height of the discs can be varied. The cut shows the single hole at the bottom of the fixed disc. When used in this position, we find that adjusting the height so that the center of the disc is always at the height of the reactor's shoulder gives fair comparisons.

The instrument is shown as connected with the 120-volt D. C. power line, with telephone receiver and two 10-watt lamps in series with stylus and plate. By using a lamp in each side of the circuit, no dangerous short-circuit can occur, although if the connections on the stylus and telephone receiver are not carefully covered, the psychologist or the reactor may receive a shock on touching the test-plate with the bare hand.

The telephone receiver is for most purposes as reliable as any available form of recording apparatus, in spite of the theoretical advantages of the latter. In cases where special study is to be made of the durations of touches, and where other considerations may be sacrificed to this, recording markers may be substituted for the telephone receiver, with appropriate change of current. For such work, a condenser must be used, of course. Dry cells may be used with the telephone receiver; but in that case, the durations of the touches can not be readily noted. 60 cycle, A. C. is even better than D. C. For most purposes, it is sufficient to count the number of touches in the standard time, and to note whether the reactor belongs to the type which readily follows instructions to move the stylus away from the brass immediately on making contact, or belongs to the type which tends to press the stylus against the brass.

The angular adjustment of the plate may be made once for all: we use it tilted upwards 15 degrees from the vertical. The vertical adjustment for each reactor, whether the tests are made in sitting or standing position, is of primary importance, if reactors are to be compared fairly. If it is desired to alternate between sitting and standing positions with a given reactor, two rings, one above and one below the clamp, may be set on the vertical rod in the proper positions, and the plates may then be slipped from the one position to the other without further measurement.

The rotation of the test-plate is also very important. By this means, each hole used is brought into exactly the same position for use, and hence in using a series of holes on one reactor, the arm and eye adjustment may be made identical for the several holes. In the older forms of plate, there are two holes which are adjacent in the series of sizes, but which lie, one at one end of the top row, and the other at the other end of the bottom row. We have found that the change from one of these holes to the other introduces a decided 'step' in the numbers of touches in the holes in running through the series. Such disturbances are entirely avoided in our test-plate.

The test-plate is six inches in diameter, and has places for 15 holes arranged in a circle of two and three eighths inch radius. In the instruments used so far, only eleven of these holes have been put in, namely: holes of 8, 9, 10, 11, 12, 14, 16, 18, 20, 22, 24, sixty-fourths of an inch diameter. This series seems excellently adapted to cover the requirements for work on children and adults; but any other series on either the inch or the millimeter scale may be put in.

The vertical rod which carries the plates is three fourths of an inch in diameter, and 30 inches long, and is mounted on a cast-iron base, designed for this instrument, which allows very convenient placing on any table.

II. Double Tapping Plate

In place of the usual single tapping plate, we are using a double plate, Fig. 2, and are requiring the reactor to strike the stylus alternately on the two plates. In this way, accurate records may be obtained by a make-and-break counter, such as the Ewald chronoscope. (*Some* Ewald chronoscopes will work accurately.) With the single plate, no form of counter yet devised will give accurate counts, and many serious errors may be made with a counter and plate which record fairly well for some reactors, and not for others. With the double plate, the necessary method of stroke gives more reliable contact, and slight oxidization of the plate is less important.

By using a definitely coördinated movement, we not only eliminate tremor-tapping, but also avoid some of the pseudo-practice effects, due to changes of method in holding the stylus, in force of blow, and in force of grip of stylus: changes which come in suddenly in many cases with the single plate.

The plates in this instrument are 3 inches square, separated by a bakelite block 1 inch wide, which rises 3/16ths of an inch above the surface of the plates. This block prevents the reactor from drawing the stylus across the plates. The plates are mounted on a cast-iron base, so heavy that it need not be fastened to the table. The base is hollow, and the condenser is to be placed under it, and connected to binding posts provided on the inside. The use of a condenser prevents the rapid deterioration of the plate surface. The Western Electric 1 microfarad condenser works well.

We are no longer able to undertake the manufacture of either of these instruments, and they will be supplied by Stoelting.

ELIMINATING THE PITFALLS IN SOLVING CORRELATION: A PRINTED CORRELATION FORM

BY HERBERT A. TOOPS

Institute of Educational Research, Teachers College

The time and labor involved in calculating the Pearson r, even when using a plotting method, is great. Then, also, where many individual product-moment multiplications, involving plus and minus signs, have to be made and summated, errors are likely to be numerous. In consequence, the research worker does not employ the method of correlation in many cases where the correlation coefficient, or the regression coefficient, would throw much light on the statistical results of his experiments.

The plotting method outlined below will greatly facilitate the computation of correlation coefficients in point of both economy of time and reduction of the chances for making errors in the computations. By use of this method, the laborious computation of individual product-moments is eliminated; negative deviation steps of other methods are replaced by steps all of which are positive, thus making all summations positive and readily done on the adding machine; there are no corrections, d_y or d_x, to be computed; and a printed table does all required extensions of table step deviations and their squares by their respective frequencies.

Fig. 1 shows in detail the general scheme of the method. This method combines in practical form four principles of use in "short-cut" calculations in solving correlation:

(*a*) When we expand the algebraical expression,[1]

$$\Sigma(X - Y)^2 = \Sigma X^2 - 2\Sigma XY + \Sigma Y^2, \qquad (1)$$

[1] Chapman, J. C., 'A Method of Calculating the Pearson Coefficient of Correlation, without the Use of Deviations or Cross Multiplying,' Psych. Bulletin, Vol. 16, No. 11, 1919, pp. 369–370.

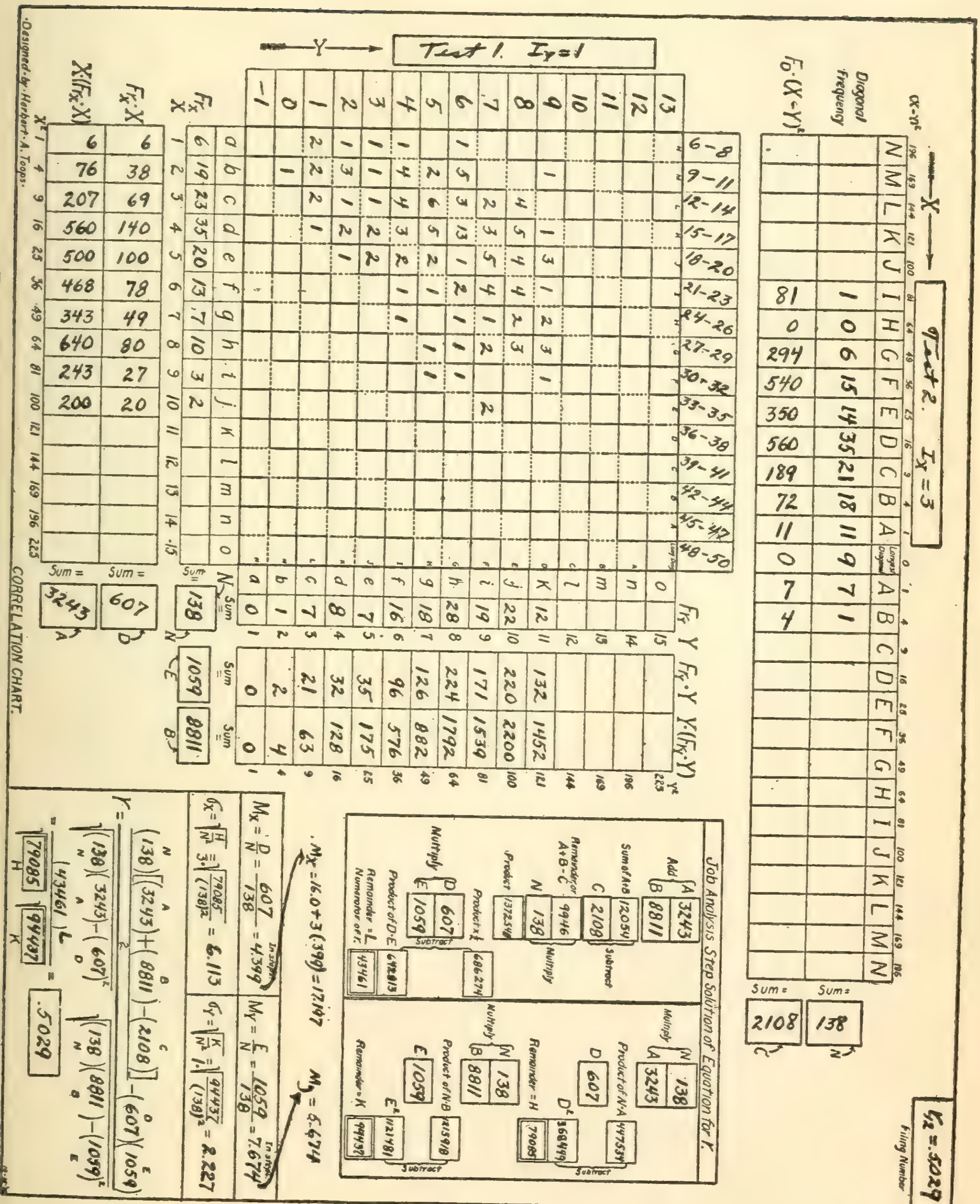


Fig. 1. The Solution of r by Use of the Printed Correlation Form.

we find that,

$$\Sigma XY = \frac{(\Sigma X^2 + \Sigma Y^2) - \Sigma(X - Y)^2}{2}. \tag{2}$$

If the right-hand member of equation (2) is substituted for ΣXY in the equation for r involving gross measures,

$$r = \frac{N \cdot \Sigma XY - \Sigma X \cdot \Sigma Y}{\sqrt{N \cdot \Sigma X^2 - (\Sigma X)^2}\ \sqrt{N \cdot \Sigma Y^2 - (\Sigma Y)^2}}, \tag{3}$$

a formula results which allows us to substitute the squared terms of equation (2) in place of the laborious product-moment sum, ΣXY.

(*b*) If the average of both distributions is assumed at the zero step, then all step deviations become positive, eliminating all negative step deviations, and the gross measure formula, (3), applies to the step deviations considered as positive gross measures.

(*c*) If X and Y be taken as table steps in a correlation plot, and the $(X - Y)$ differences be plotted by compartments, it will be found that for *any given diagonal row* the $(X - Y)$ differences are all alike; this allows one to *count* the *diagonal frequencies* by rows and then multiply the total diagonal frequency of a given diagonal row by its proper $(X - Y)^2$ to obtain that diagonal row's contribution to $\Sigma(X - Y)^2$ of formula (2). Thus, the simple procedure of counting diagonal frequencies replaces the laborious product-moment computations of other methods.

(*d*) The first power of the table steps, and also the squared terms employed in formula (2), each multiplied by all possible frequencies up to as high as 50, are readily grouped in tabular form on a printed $8\frac{1}{2}$ x 11-inch chart, the use of which, in making extensions of table steps and their squares by frequencies, eliminates the larger part of the labor involved in making such extensions by the usual methods. The adding machine performs all the positive summations and affords a check when comparing the original entries with the record which the listing-adding machine supplies.

The bearing of these four facts upon the method will be suggested by the formula itself, which is derived by substituting equation (2) in equation (3):

$$r = \frac{\frac{N}{2}[(\Sigma X^2 + \Sigma Y^2) - \Sigma(X - Y)^2] - \Sigma X \cdot \Sigma Y}{\sqrt{N \cdot (\Sigma X^2) - (\Sigma X)^2}\,\sqrt{N \cdot (\Sigma Y^2) - (\Sigma Y)^2}}. \tag{4}$$

In this formula, X and Y are graph steps, or table deviations; ΣX and ΣY are arithmetical sums of positive quantities. The corrections for an assumed mean become an integral part of the formula and do not have to be calculated separately as in other plotting methods. The formula is quite general and applies to the actual gross measures as well as to the table steps considered as gross measures.

That the $(X - Y)$ differences for any given diagonal row is a constant for every compartment in that diagonal row may be easily proven mathematically. This is shown schematically also in Fig. 2. The equation of the longest diagonal row of compartments is, $Y_0 = X_0$; of the first diagonal row downwards, $Y_1 = (X_1 - 1)$, since this row differs from the preceding diagonal row only by its y-intercept, a constant quantity for the row considered. Generalizing, the equation of the nth diagonal row downwards is, $Y_n = (X_n - n)$. Now for this nth diagonal row, the $(X - Y)$ difference is, $X_n - Y_n = X_n - (X_n - n) = n$, a constant, the ordinal number of the diagonal row below the longest, or main diagonal row. For diagonal rows above the longest diagonal, n is negative in sign but it is squared in formula (2) and so becomes positive in its summation.

Referring to Fig. 1 for illustrations, the itemized procedure in calculating r is as follows:

1. Enter the names of the X- and Y-variables in the proper boxes on the correlation sheet.

2. Enter the classes of the X-variable in ascending order

etc.

Y-steps					
5	−4 *D*	−3 *C*	−2 *B*	−1 *A*	0
4	−3 *C*	−2 *B*	−1 *A*	0	1 *A*
3	−2 *B*	−1 *A*	0	1 *A*	2 *B*
2	−1 *A*	0	1 *A*	2 *B*	3 *C*
1	0	1 *A*	2 *B*	3 *C*	4 *D* *D*
	1	2	3	4	5

etc.

X-steps

Fig. 2. Schematic diagram showing how the $(X - Y)$ difference for any given diagonal row is a constant, thus allowing the counting of diagonal frequencies to take the place of product-moment multiplications. The quantity in each compartment is the *X*-step of that compartment minus its *Y*-step. Since the $(X - Y)$ differences are squared, thus becoming positive, before being used in formula (2), there are no negative signs to be taken into account. The frequencies of a given diagonal row, identified by a small capital letter in Fig. 1, are counted and then multiplied by the positive $(X - Y)^2$ quantity of the diagonal row considered. For the longest diagonal row, $(X - Y)^2$ is zero; for each of the diagonal rows just above and below, each lettered *A*, $(X - Y)^2$ is 1; for the two rows lettered *B*, $(X - Y)^2$ is 4; for the two rows lettered *C*, $(X - Y)^2$ is 9, and so on.

from left to right. Use such class intervals, I_x, that the maximum number of classes will result.[1]

[1] In order to find the class intervals, I_X and I_Y: In the variable under consideration, subtract the lowest gross score obtained from the highest gross score obtained; to the remainder add 1; divide the quantity thus obtained by the maximum number of classes of the printed chart, here 15; the quotient, raised to the *next highest* integer is the correct maximum class interval. Example: The lowest score on a test is 13; the highest, 40. Then the class interval, I_X, is:

$$I_X = \frac{(40 - 13) + 1}{15} = \frac{28}{15},$$

which, when raised to the next highest integer, is 2.

As a desirable convention the first step class should always *begin with a multiple of the class interval;* thus, in the above example, the classes would run (class interval = 2):

Class	Step
12–13	1
14–15	2
16–17	3
Etc.	

3. Enter the classes of the Y-variable in ascending order from bottom of the sheet to top.

4. Plot the paired measures as tally marks in the compartments using the classes determined in procedures 2 and 3 as coördinates. When completed, enter the frequencies of each compartment as a small figure in the lower left hand corner of each compartment. (To avoid confusion the tally marks are omitted in Fig. 1 and these compartment frequencies only are shown as a figure in the center of each compartment.)

5. Add vertical frequencies by columns, entering column frequencies at the feet of the proper columns in the row, $Fr._X$.

6. Add horizontal frequencies by rows, entering the row frequencies at the right hand ends of the proper rows in the column, $Fr._Y$.

7. Add the frequencies of the several diagonal rows separately, guided by the dotted lines, and enter the several diagonal row frequencies in the correspondingly lettered spaces provided at the top of the page in the horizontal row, marked, "Diagonal Frequency." (A small capital letter serves to identify the diagonal row; a given diagonal row's total frequency is to be entered under the corresponding large capital letter. Note that A means that diagonal row which has an $(X - Y)$ value of ± 1; B, an $(X - Y)$ value of ± 2, and so on. Note also that, on this chart, the N-diagonal frequency is the sum of the frequencies of but one compartment, M of two, L of three, and so on. A celluloid stencil or cardboard serrated edge may be helpful in relieving eye strain when counting the diagonal frequencies, and will contribute to accuracy in the counting.)

8. Add all vertical, horizontal, and diagonal marginal frequencies. Each sum must add up to N, the total number of cases, in the respective N-windows.

9. Multiply the $Fr._X$ frequencies by the respective X table steps obtaining the row of products, $Fr._X \cdot X$; likewise multiply the $Fr._Y$ frequencies by the respective Y table steps obtaining the column of products, $Fr._Y \cdot Y$.

10. Multiply the X steps respectively by the $Fr._X \cdot X$ products above obtained thus securing the $X \cdot (Fr._X \cdot X)$ products; likewise, multiply the Y steps respectively by the $Fr._Y \cdot Y$ products above obtained thus securing the $Y \cdot (Fr._Y \cdot Y)$ products.

11. Multiply the several diagonal frequencies by their respective $(X - Y)^2$ values, thus obtaining the $F_D \cdot (X - Y)^2$ products.

12. On the adding machine add the products separately of each of the five rows and columns of procedures 9, 10 and 11. In the example, Fig. 1,

$$\begin{aligned} \Sigma Fr._X \cdot X &= \Sigma X &&= D = 607 \\ \Sigma Fr._Y \cdot Y &= \Sigma Y &&= E = 1059 \\ \Sigma X \cdot (Fr._X \cdot X) &= \Sigma X^2 &&= A = 3243 \\ \Sigma Y \cdot (Fr._Y \cdot Y) &= \Sigma Y^2 &&= B = 8811 \\ \Sigma F_D \cdot (X - Y)^2 &= \Sigma (X - Y)^2 &&= C = 2108 \\ & && N = 138. \end{aligned}$$

13. Substitute the values thus found in formula (4) and solve for r. (To facilitate this operation, a skeleton formula is employed. Since the chart is a printed form chart the required six sums always appear in a definite place, the lettered windows as given above, in each case. By substitution of the above-lettered notation for the algebraic notation of formula (4), the skeleton formula in the lower right-hand corner of Fig. 1 results. The substitution of the letter-named window sums in the skeleton formula is a process understood by clerical workers who have no training in the statistical notation which the letters replace.)

It will be readily apparent that Table I., which is complete for only the frequencies 1 to 9 inclusive, in its complete form may be used to do all the extensions of marginal frequencies by the table steps and their squares, and will replace all of the work of procedures 9, 10 and 11 above. Table I. (rows a, b, c, $\cdots$, and rows, a^2, b^2, c^2, $\cdots$) gives for the first nine frequencies, or 1 to 9, the $(X \cdot Fr._X)$, $(X^2 \cdot Fr._X)$, $(Y \cdot Fr._Y)$, and $(Y^2 \cdot Fr._Y)$ values, and also (in rows A, B, C, $\cdots$), the

$(X - Y)^2 \cdot F_D$ values. The author makes use of a larger printed table, $8\frac{1}{2}$ x 11 inches in size, which gives the products by units for all frequencies from 1 up to 50, and by tens from thence on to 100. With two workers,—one to read frequencies and work the adding machine, the other to read the Table I. entries corresponding to the marginal frequencies—the required quantities may be taken from the table almost as fast as they can be placed into the adding machine, since just one figure is taken from each row of the table in succession down the page. The extension of table steps and their squares by frequencies may also be found by using the multipliers printed directly on the chart, as well as from Table I. which uses the a and a^2, b and b^2 letters as guides to locate the table entries in turn. The products, however obtained, may be either written down in the proper spaces of the correlation sheet, or be added directly on the adding machine to obtain the respective sums. In the latter case the adding machine record serves as a check on the accuracy of the original extensions.

A calculating machine is very efficient in reducing the skeleton formula to the final form involving the simplified reductions, L, H and K. This reduced form may then be easily solved by slide rule. The integral quantities, H and K, enter directly into the computation of the standard deviations.

In all operations, two workers are preferable: A, to enter tally marks of paired scores read to him from the original data by B; later, A to read in succession to B the letters and corresponding marginal frequencies, the tabulated products of which are found in Table I. by B; then, finally, A to enter these products on the printed graph sheet while B enters them into the adding machine. Checking may be readily accomplished by A and B interchanging places and checking the products with the printed listed record of the adding machine.

The solution of equation (4) may be simplified, for the benefit of persons not accustomed to the solution of equations,

Table I.

Table of First and Second Powers of Steps by Frequencies

Fr.		1	2	3	4	5	6	7	8	9		Fr.
	a	1	2	3	4	5	6	7	8	9	a	
A	a^2	1	2	3	4	5	6	7	8	9	a^2	*A*
	b	2	4	6	8	10	12	14	16	18	b	
B	b^2	4	8	12	16	20	24	28	32	36	b^2	*B*
	c	3	6	9	12	15	18	21	24	27	c	
C	c^2	9	18	27	36	45	54	63	72	81	c^2	*C*
	d	4	8	12	16	20	24	28	32	36	d	
D	d^2	16	32	48	64	80	96	112	128	144	d^2	*D*
	e	5	10	15	20	25	30	35	40	45	e	
E	e^2	25	50	75	100	125	150	175	200	225	e^2	*E*
	f	6	12	18	24	30	36	42	48	54	f	
F	f^2	36	72	108	144	180	216	252	288	324	f^2	*F*
	g	7	14	21	28	35	42	49	56	63	g	
G	g^2	49	98	147	196	245	294	343	392	441	g^2	*G*
	h	8	16	24	32	40	48	56	64	72	h	
H	h^2	64	128	192	256	320	384	448	512	576	h^2	*H*
	i	9	18	27	36	45	54	63	72	81	i	
I	i^2	81	162	243	324	405	486	567	648	729	i^2	*I*
	j	10	20	30	40	50	60	70	80	90	j	
J	j^2	100	200	300	400	500	600	700	800	900	j^2	*J*
	k	11	22	33	44	55	66	77	88	99	k	
K	k^2	121	242	363	484	605	726	847	968	1089	k^2	*K*
	l	12	24	36	48	60	72	84	96	108	l	
L	l^2	144	288	432	576	720	864	1008	1152	1296	l^2	*L*
	m	13	26	39	52	65	78	91	104	117	m	
M	m^2	169	338	507	676	845	1014	1183	1352	1521	m^2	*M*
	n	14	28	42	56	70	84	98	112	126	n	
N	n^2	196	392	588	784	980	1176	1372	1568	1764	n^2	*N*
	o	15	30	45	60	75	90	105	120	135	o	
O	o^2	225	450	675	900	1125	1350	1575	1800	2025	o^2	*O*
	p	16	32	48	64	80	96	112	128	144	p	
P	p^2	256	512	768	1024	1280	1536	1792	2048	2304	p^2	*P*
	q	17	34	51	68	85	102	119	136	153	q	
Q	q^2	289	578	867	1156	1445	1734	2023	2312	2601	q^2	*Q*
	r	18	36	54	72	90	108	126	144	162	r	
R	r^2	324	648	972	1296	1620	1944	2268	2592	2916	r^2	*R*
	s	19	38	57	76	95	114	133	152	171	s	
S	s^2	361	722	1083	1444	1805	2166	2527	2888	3249	s^2	*S*
Fr.		1	2	3	4	5	6	7	8	9		Fr.

by use of the Job Analysis form. In using this form, the labeled windows in the left-hand column of each of the two sections are to be filled out first by merely copying into them the sums already determined and entered in the correspondingly labeled windows. Thereupon, the reduction of the formula is merely a matter of blindly following out the directions of, "add," "subtract," "multiply," etc., writing the sums, remainders, and products in the designated windows in turn. The three quantities, L, H and K, enclosed by a double rectangle, are the quantities which, when solved by slide rule in the second reduction of the skeleton formula, yield r. The slide rule is so efficient in this last step that it is not advisable to make out job analysis steps involving extraction of the square root.

In summary, this method of solving r simply is: Find vertical, horizontal and diagonal marginal frequencies, adding the corresponding lettered table products found in Table I. to obtain the lettered window sums which enter into the skeleton formula.

To find the average of the X-variable, or M_X: substitute the proper window quantities in the skeleton formula for M_X and solve; *the answer is in terms of table steps.* The resulting table steps must be interpolated to find the value of the average in terms of the X-variable scores. Thus, in Fig. 1, $M_X = 4.399$ table steps. Now the *fourth* table step in the X-variable has the *face-value* of 16.0; and .399 of a table step, where the class interval, $I_x = 3$, is $.399 \times 3 = 1.197$. Or, the true average of the X-variable is $16.0 + 1.197 = 17.197$. Similary, the average of the Y-variable is 7.674 table steps of the Y-variable, which interpolated gives the true average of the Y-variable, which is 5.674, $(I_y = 1)$.

The formula for the standard deviation is:

$$\sigma_X = I_X \sqrt{\frac{N\Sigma X^2 - (\Sigma X)^2}{N^2}} = I_X \sqrt{\frac{H}{N^2}}. \tag{5}$$

Where, I_X is the class interval of the X-variable. Note that H is found in the reduction of the formula for r and need not be recalculated. The standard deviation of the Y-variable is found by substituting Y for X in formula (5). The magnitudes of I_X and I_Y, the class intervals, do not enter into the computation of r; they are always to be made as small as the data will permit in order to reduce the grouping effect which occurs whenever data are grouped into coarser classes.

In practice it has been found that the only operation at all likely to result in errors is the counting of the diagonal frequencies. This may be partially obviated by use of a serrated edge. A very effective check, not only of the counting of the diagonal frequencies but also of the numerical work of the whole table, lies in the fact that the sum of the quantities $(A + B - C)$, when simplified in the skeleton formula, *must be an even number*. This follows from a consideration of the fact that ΣXY in equation (3) is always an integral number; therefore, $\frac{1}{2}(A + B - C)$ must be an integral number, but this is only possible when $(A + B - C)$ is an even number. If then a single error is made in either ΣX^2, ΣY^2 or $\Sigma(X - Y)^2$, it will very likely make $(A + B - C)$ an odd number, which odd number would always be a certain indication that an error had been made. The only safe method, of course, is to recheck every operation as is customary in all statistical work. With this method, intelligent clerical workers have been found capable of solving correlations with as much accuracy as trained statisticians. The total time required to solve a correlation is greatly reduced over that required by previous methods; while, with a single checking of all operations, errors in the final r almost never occur. Since the reduced numerator, L, and the quantities, H and K, are always integral numbers, it follows that the final value of r is never influenced by the extent to which the previous decimals, used in other methods, may have been carried out. The final operation is the only one in which decimals enter.

The speed of the plotting, and also its accuracy, may be increased by the use of the plotting device described below, to be used with the printed correlation sheets described above.

A Correlation Plotting Device.—In efficiency, the combination of printed sheet and plotting machine has been found through extensive try-out to yield Pearson correlation coefficients in from one-third to one-fourth of the time taken by other methods. In point of speed, it has been found possible to plot consistently twenty paired measures per minute with this device. The principle is quite general and will be useful in all cases where it may be required to locate quickly any desired compartment of a contingency table by means of its coördinates, as in reading products in the body of a large multiplication table where the two factors are printed as coördinates along the two edges.

The essential parts of this machine are a framework of wood, *A*, bearing a transparent celluloid stationary runner, *B;* a device for holding coördinate paper ($\frac{1}{4}$-inch squares) represented in the picture by the knobs, *C* and *D*; and a movable slide, *E*, on which the printed plotting paper may be fastened with thumb tacks.

In use, the printed correlation paper is tacked to the slide, *E*. A strip of $\frac{1}{4}$-inch coördinate paper is fastened to the edge of the stationary runner by inserting the tips of the coördinate paper strip under the washers found under *C* and *D* and tightening these knobs. The classes of the *Y*-variable are entered on this strip, while the classes of the *X*-variable are entered along the top of the printed sheet in their usual space. The operator manipulates the machine while the teller reads to him the paired measures, reading first the *X*-measure and then the *Y*-measure. The operator holds the slide, *E*, with his left hand. As the teller calls out to him, say 16, the operator moves the slide and with it the printed sheet until the 16-column of the *X*-variable is the first column visible to the right of the edge of the celluloid runner and paper slip, as in Fig. 3. He does this almost instantly by watching only the *X*-classes as they disappear beneath the paper strip, so that, by the time the teller has called the value of the second or *Y*-measure say 8, the operator is in position with his right hand to make a pencil mark tally along the edge of the paper strip in that compartment which is directly beside 8 of the paper strip. This device is thus one for bearing the marginal frequencies of the *Y*-variable into the body of the table to secure accurate and speedy plotting; its psychological advantage being that the operator need give attention to but one of the coördinates at a time. There is almost no likelihood of error in the plotting if the operator calls back to the teller the two measures as heard and is checked by the teller.

By use of the serrated heavily inked-in diagonal boxes on the celluloid runner at *F*, it is easy to count the frequencies in any diagonal row without confusing them with the frequencies of any other diagonal row. Error in entering the diagonal sums in the "Diagonal Frequency" row is made mechanically impossible by having accessible only the one cut-out window, *G*, of that diagonal row whose frequency is being counted by the aid of the serrated boxes. This window, *G*, is so aligned with the last square at the top of the diagonal row as to be open only above the proper entry space of the diagonal row being counted.[1]

[1] The correlation sheets and table of products of deviation steps and their squares by frequencies may be obtained from the author, Institute of Educational Research, Teachers College. The correlation sheets have recently been revised as a result of two years' daily use in research bureau grading and testing work. After plotting, fifteen minutes is more than ample time for the solution of a correlation such as that shown in Fig. 1.

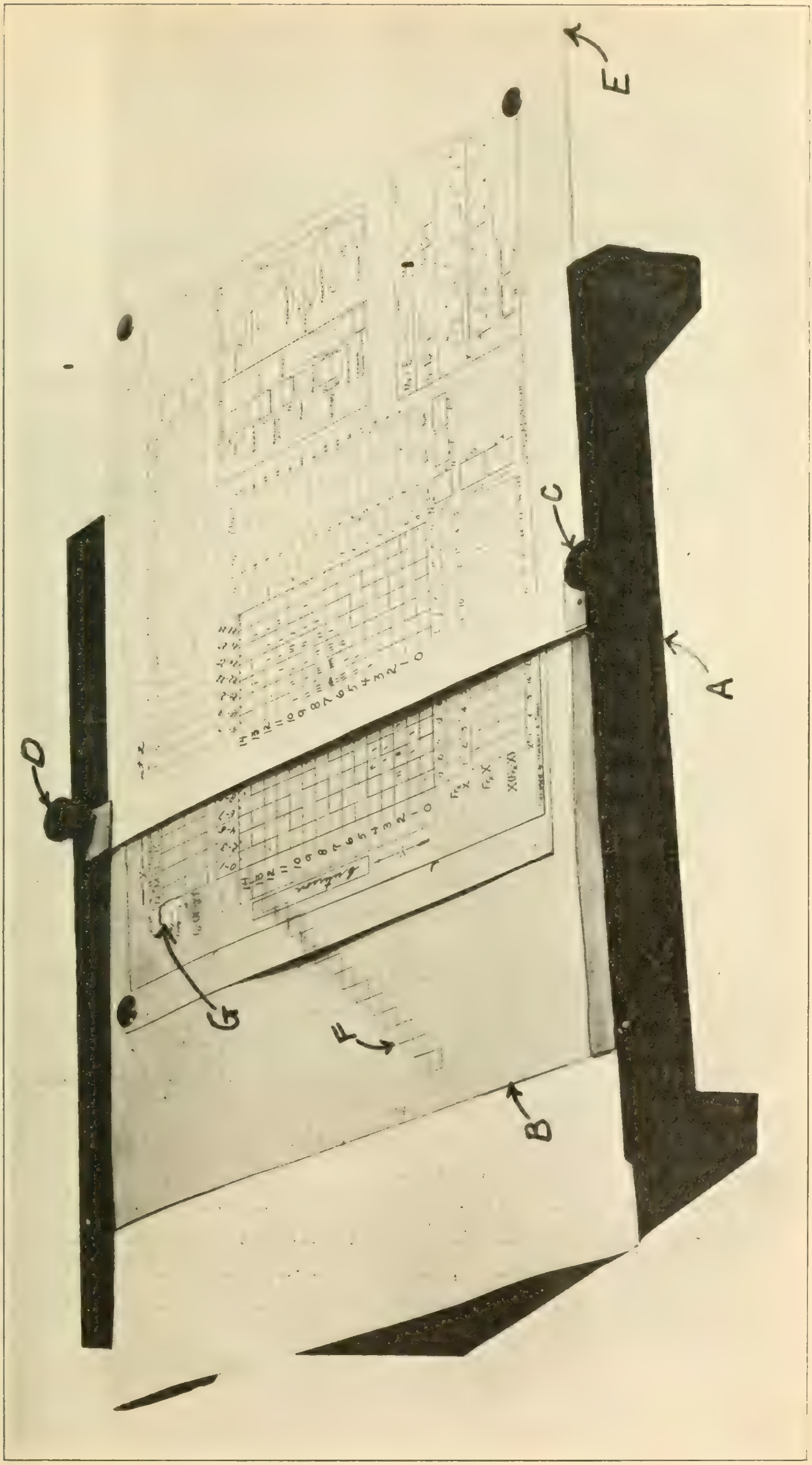


Fig. 3. A correlation plotting device

THE RÔLE OF SYNÆSTHESIA IN LEARNING

BY R. H. WHEELER AND T. D. CUTSFORTH

University of Oregon

Introduction.—A preliminary analysis of a case of synæsthesia by means of the introspective method clearly indicated that it was worth while to undertake several intensive studies of the phenomena first, for the reason that detailed introspective descriptions even from one trained synæsthetic reagent seemed to throw considerable light upon the nature and functioning of synæsthesis itself and upon several systematic problems the status of which seems to be in doubt as far as experimental results are concerned. Secondly, since our synæsthetic reagent is blind and since we were able to check his introspections against a second blind reagent who is asynæsthetic,[1] it seemed worth while to contribute indirectly to the literature on the psychology of the blind.

The purpose of this particular investigation was to ascertain the value and functioning of synæsthetic processes in learning and to draw such conclusions as might be possible regarding the nature of synæsthesis itself from a functional point of view.

Method, Apparatus and Observers.—Observer *A* is a graduate student at the University of Oregon with four years of introspective training. Observer *B* is likewise a trained introspector and was a major student of advanced standing at the time his introspections were obtained. *A* lost his sight at the age of eleven and *B* at the age of nine. *B* was employed as a 'check' reagent.

The material for this investigation consisted of two sets of nonsense syllables, one for tactual presentation and one for auditory presentation. Each set contained four series of

[1] The term 'asynæsthetic' was suggested to the senior author by Dr. Edwin G. Boring to designate a person's mental equipment which is not synæsthetic.

ten syllables each. The syllables for tactual presentation were punched by a Braille machine on rectangular pieces of tag-board 5 × 10 centimeters in size. These cards were placed on the table in front of the reagent. After he inspected the first card tactual fashion it was removed by the experimenter while the reagent lifted his hand and arm for an interval of three seconds. By the end of this interval a new card had been substituted and as the reagent lowered his hand he found his fingers in the correct position to begin reading the next syllable. An interval of six seconds was given between series presentations. A series was considered learned when the reagent was able to anticipate the syllables correctly. After the learning, the reagent recited the syllables, spelling each one, and then gave as complete an introspection as possible upon the learning and upon the immediate recall. Twenty-four hours after the original learning the reagent was asked to recall the syllables and to introspect on this delayed recall.

In the auditory presentations the experimenter read the syllables as distinctly as possible, without spelling the letters, and at three-second intervals, allowing six seconds between series presentations. Introspections were obtained after an immediate and after a twenty-four-hour delayed recall.

The learning was done throughout under optimal laboratory conditions. A second series of syllables was not learned until the preceding series had been recalled for the last time. Our rather slow presentation was a feature of experimentation introduced in order to enable the reagent to give more detailed introspective descriptions of his learning than would have been possible if he had felt hurried. We were not interested in the learning process *per se*.

Typical Introspective Data

Our introspections were so long and detailed, some of which covered four or five typewritten pages, that it is here impossible to present more than one complete introspection from each reagent; but the reader will be able to grasp something of the great detail into which the observers carried their introspective descriptions and will readily understand why it has been possible to obtain results of a fairly elaborate character.

1. *Observer B.* [Parentheses indicate the reagent's interpretations of his own mental processes after they had taken place. Brackets indicate comments inserted by the editors. Each introspection is numbered for the purpose of reference.]

"Was aware of the task in terms of kinæsthetic tensions centered about my eyes, brows, shoulders, left arm and hand, together with focal auditory perceptions of the experimenter's instructions and incipient tactual and motor imagery of masses of points on a dimly localized tag-board beneath my fingers. For an instant there was an entire exclusion of other processes. (These experiences meant to me both an attitude of acceptance and of expectation.) As the series was presented I was first conscious of blunt tactual sensations and of finger movements. Then these vague points became grouped into spacial arrangements which latter constituted my tactual perceptions of the letters. As the grouping took place I was less conscious of the individual points, and became aware of a broadening of the scope of tactual attention in which all of the points of a given letter stood out with approximately equal distinctness. At the final stage of a perception that part of a letter at which a point is missing stood out in focal consciousness, or the outstanding feature was a 'bar' made by two or more points in a line and extending in a certain direction—horizontal, vertical, oblique down and to the left, or what not. At times the length of the bar was the focal feature here, and at other times the direction of the bar was the focal feature. In still other instances the outstanding feature of the perception finally became an angle formed by a group of points. For example my perception of 'c' in one of the syllables consisted of a right angle which opened to the left and whose apex was pointing to the right. In 'd' the right angle opens to the left but the apex points upward and to the right. These features stood out in terms of tactual and kinæsthetic imagery. The awareness of an angle consisted first of a combined tactual-kinæsthetic perception of the lower end or side; the perception always began at the bottom and developed upward and to the right in the direction of finger movement. There then developed incipient finger and arm movements of tracing the right angle. In letters like 'm' and 'p,' which if solid would make a similar angle kinæsthetic imagery, appeared with the movement longer on one side. The perception of each letter also involved incipient tactual-motor imagery of finger and arm movements having to do with punching these letters by the hand method of stylus and slate. This imagery was always more distinct in such letters as did not appear in the context for some little time."

"During the process of learning each letter lingered in consciousness momentarily, followed by vocal-motor imagery. As the vocal-motor imagery developed in each instance the tactual-motor features began to disappear. Consciousness of any syllable as a whole was entirely vocal-motor-auditory. During the intervals between presentations I was engaged in repeating the preceding syllables vocal-motor fashion. I noticed that in retaining the syllables the last feature of the tactual-kinæsthetic perception to disappear was invariably a single, dominating, incipient finger and arm movement. As the presentations continued I found myself trying to anticipate the syllables, using auditory-vocal-motor imagery wherever possible. This method proved more efficient because tactual-kinæsthetic imagery involved the details of three letters while the former imagery involved a simple sound or incipient throat movement. Only when I failed to anticipate a syllable in vocal-motor fashion did I find myself resorting to the slower method of using tactual-motor imagery. I learned the first and last syllables at the outset and the middle ones last."

"During the immediate recall I was first aware of the 'bih' in terms of auditory imagery in rising inflection as if I were about to ask the question 'How do you pronounce it?' As the vocal-motor-auditory imagery developed there also appeared rather definite tactual-motor imagery of the syllable in which latter imagery the individual letters stood out prominently but only for a brief instant. There then occurred a schematic grouping of the syllables in terms of incipient movements of my shoulders as if throwing the subject matter of the series into separated spaces; and with each rhythmic emphasis in the motor tendencies I had the bare beginnings of auditory imagery of the syllables themselves. This schematic grouping arranged the first four syllables into one group, the next four into a second group and the remaining two in a third group. In anticipating the syllables, vocal-motor-auditory fashion, this schematic kinæsthesis was intensified at the beginning of each group. The second, third, and fourth syllables appeared in vocal-motor imagery, with no tactual-motor tendencies to spell out the letters until I spelled them for the experimenter. Fleeting and schematic imagery of letters came in tactual-motor fashion wherever there was difficulty in recalling the syllables." [Six presentations were required for the learning and the observer recalled the syllables correctly save for one letter in one syllable.]

2. Delayed recall: "First there appeared a diffused motor consciousness of strains about the brows, throat and chest; motor and tactual imagery appeared of my arms outstretched before me and of making movements of investigating Braille syllables. The first syllable came readily in vocal-motor imagery (suggested by the tactual-motor attitude just described). There was a slight tendency to translate the 'b' of 'bih' into tactual-motor imagery. The three syllables which followed came in auditory-vocal-motor imagery and with very little effort. As I recalled the fourth there first came a schematic visual-kinæsthetic image of a printed 'L'; no paper was visible in this image and no background save the 'nothingness' which is always associated with my visual imagery. The fifth syllable was hard to recall. At first I had 'vez' in auditory imagery which I found myself at once correcting to 'veh,' tactual-motor fashion. The antecedent of this correction consisted of developing tensions about the forehead

and trunk which grew more intense as my attention began to shift toward the act of anticipating the next syllable. There was kinæsthetic imagery of jerking my arm back; breathing seemed to be checked momentarily, with tensed diaphragm and a vigorous vocal-motor-auditory: 'No.' The next syllable, 'pog,' appeared in auditory vocal-motor imagery, together with incipient tendencies to vocalize it before I was certain that it was correct. Certainty here seemed to consist of the persistence of the auditory-vocal-motor imagery of the syllable, together with a developing tactual-motor image of the three letters and a tendency to smile. Then I paused; sudden tensions developed about the forehead, trunk, and especially in the jaws and throat; there came fleeting auditory imagery: 'What next?' I had tendencies to move my arm and hand as if feeling of letters (this together with the tensions constituted an effort to recall the next syllable.) Then I had tactual imagery of the laboratory table, cards, chair, and auditory imagery of the experimenter's voice. (These latter constituted a consciousness of the task.) Then I found myself exclaiming vocal-motor fashion, 'Oh,' as 'double U' appeared in vocal-motor-auditory imagery. Then the syllable 'nuz' suddenly appeared and I at once found myself vocalizing it. The remainder of the series came easily and in auditory-vocal-motor imagery."

3. *Observer A.* [Syllables were dop, jan, vum, qoh, qez, pif, xap, loy, pef, hov and were thus arranged in a series for the purpose of watching the behavior of synæsthetic imagery.]

"I first went through the series with the intent of becoming generally familiar with the syllables, and of constructing a pronunciation for them or a name for each. I grouped the first three together in terms of vocal-motor-auditory imagery, thus, 'dop-jan-vum.' I paid no attention to the tactual-kinæsthetic elements in the learning or in the perceptions of the letters. As fast as I inspected each syllable the appropriate synæsthetic visual image appeared at my finger tips, the coloring and brightness of the imagery being determined by the letter. When I then found myself vocalizing the syllable, the syllable itself tended to take on the color of the initial letter. As my attention shifted from the procedure of inspecting the letters, tactual-kinæsthetic fashion, the visual imagery which at all times remained dominant in consciousness shifted to the center of my field of vision. Throughout, the tactual-motor processes were dim, vague, indescribable experiences of which I know nothing except in terms of their visual associates. [In other words the former were cues for the appearance of the latter.] It was not until the third repetition that I was able to anticipate any more than the first three syllables. My learning procedure consisted chiefly of visualizing the color for each syllable in a schema. The first three syllables assumed their place in a straight line to the left in the schema, and along with this there took place auditory-vocal-motor imagery of them. The remaining syllables were learned when I could fix a place for them in this visual schema. The fourth syllable to be thus placed in the schema was the eighth, whose color was brighter than its neighbors; this syllable occupied an elevated position in the form and considerably to the right of the first three. Between them and to the left of the eighth were spaces of neutral gray into which I later fitted the remaining syllables. After the first four or five presentations I changed my *Aufgabe* from that of retaining to that of anticipating the syllables and my method, as before, was almost entirely visual."

"As each syllable was inspected tactual-motor fashion my attention was centered visually upon the tips of my fingers; with a certain amount of finger-movement colors began to appear at my finger tips; these colors were my awareness of the letters as such; just as soon as these colors began to develop I shifted my line of regard to the center of a visual field extended out before me in space and there tried to visualize the syllable in color."

"When I centered my attention upon one section of the form, it only was clearly 'seen'; in indirect vision I could 'see' adjacent sections. The brighter or more persisting colors of these sections tended to draw attention away from the section upon which I was focusing my attention at the time. This often resulted in confusion. In the process of trying to anticipate the syllables it frequently happened that more than one color would thus attract my attention before I succeeded in clarifying a dim section in the form. I also found myself confusing colors for individual letters with colors for entire syllables. I would then have to wait for the next tactual-motor presentation in order to ascertain which color stood for a letter and which for a syllable. I had a great deal of trouble with 'pif.' The color for this syllable persistently appeared out of order."

"As the form developed it became stable and the sections standing for individual syllables were definitely differentiated. The first three syllables formed a straight line of three colors to the left; the 4th, 5th and 6th assumed a position, in colors, along a downward slant toward the right, beginning where the third syllable left off. The 7th was learned last and came to occupy a dip in the form below the others; the 8th appeared as a high crest to the right of the dip and the 9th and 10th assumed a slanting position at the right end."

"I announced that the learning was complete as soon as I was able to follow along this visual form just ahead of the tactual-motor presentation of the syllables. I fixated each section of the form and translated the color vocal-motor fashion. It was then verified by a brightening process which the colors underwent when I inspected the syllables with my fingers."

"In the immediate recall I found my visual attention centered upon the left end of this visual schema. The appearance of the first three syllables in a straight line meant to me that I had learned these together and with ease; they were dominantly grayish-green in color, shading toward a lighter tint at the extreme left; the lighter left end told me at once that the syllable began with 'd.' The visualized 'd' suggested at once the syllable 'dop,' which in turn suggested the next two syllables with no necessity of translating, consciously, from the visual schema. But as I vocalized these syllables the colors of this section of the schema brightened. I then found myself fixating the downward slant of the form; the first section of this region was colored by the letter 'q'; the second was colored by the 'z' sound in the syllable, and the third was colored by the 'x.' I fixated my attention first upon one of these colors and then upon the other. Then I translated these colors into their appropriate letter-sounds and subsequently, by means of a trial and error method, pronouncing different sounds to myself, I built up the syllables, 'qoh,' 'qez' and 'xap.' The final recall came in auditory-verbal imagery together with a brightening of the colors in the form. In turn, the 'pif' and 'loy' stood out in their own distinct colors; then at the last, 'pef' and 'hov.' In no instance was a syllable recalled first in vocal-motor fashion. Each recall was constructed rather laboriously from the verbal suggestions which came from the localized colors of the form."

4. Delayed recall: "The intent to recall the syllables of the day before led at once to colored visual imagery of the form before described. I found myself fixating the extreme left end and then moved my line of regard to the right over the first three sections which formed a straight line. This triple section was a dark, grayish-green, but lighter at the left end. I found myself saying 'd' in vocal-motor imagery as I fixated the lighter portion at the left and then, suddenly, there came the vocal-motor-auditory imagery of 'dop,' 'jan' and 'vum.' Next, moving my line of regard downward over the oblique slant which marked the next large section in the form, I noticed that the first letter of each of the next two syllables was 'q,' because the color of the form was a dark, brownish-green. As I vocalized the 'q' there at once followed the vocal-motor 'ez' together with the colors for 'e' and 'z' which appeared in the lower half of my field of vision and floated upward, fitting into their appropriate places in the form. I then found myself repeating 'q' again in vocal-motor imagery; then I had 'o' both in vocal-motor and visual imagery, the color for the 'o' taking its place next to

the 'q' in the form. Next the syllable 'pif' came easily for I could see all three letters standing out in terms of their colors. I then saw the 7th syllable as a very dark, brown section in the form but I could get no verbal imagery from looking at this color. Next my visual attention was claimed by 'loy' which stood out as the high crest or shoulder in the form; this section was dominated by the smoky blue for 'l.'"

"I now have the missing syllable—it is 'xap.' The first hint of it came in terms of the light straw yellow of the letter 'a'; then came the poorly saturated reddish-brown, followed by the 'p' in simultaneous vocal-motor-visual imagery. These colors appeared in their proper sections in the form."

"I next turned my attention to the right end of the form where I saw distinctly the colors for 'pef' and for another syllable which I could not then translate. After a few seconds the syllable 'hoy' came to me, later corrected to 'hov.' In the last section of the form the dominant color stood for 'o.' For some time I fixated that small patch of color in the form, moving my line of regard over the neutral gray surroundings, hoping other colors would appear, when I found myself vocalizing 'v,' which in turn suggested the 'h.' These latter came in combined vocal-motor-auditory and visual imagery."

"(In my recall I at first reversed the order of 'qoh' and 'qez.' This was due to the fact that both syllables began with the same letter and looked just alike in the form, for in these two instances the colors of the section were determined by the initial letter of the syllable. There was no criterion by which I could at first tell which syllable was first and which second. I have no such difficulty in cases where the three letters stand out in colors in the form.)"

Summary of Data from Observer *B*

B's procedure in the learning consisted first of perceiving the individual letters of each syllable; secondly of grouping these letters into a word, vocal-motor and vocal-motor-auditory fashion; thirdly, of repeating these syllables over and over again, holding them in consciousness as best he could in this fashion while he inspected two or more subsequent syllables; fourth, of anticipating the syllables in terms of vocal-motor imagery.

It is interesting to note that the dominating features of *B*'s tactual perceptions were invariably (1) motor imagery of angles or (2) motor imagery of punching the point letters (3) verbal imagery. These processes terminated any given act of perceiving and identified the letter.

As learning progressed, *B* tended to rely more and more upon vocal-motor-auditory imagery, resorting to tactual-motor imagery only when in difficulty. Tactual-motor imagery always turned out to be a laborious method of retaining the syllables for the reason that each letter had to be represented. By means of a motor schema, involving incipient shoulder movements, *B* grouped the ten syllables into 4, 4, and 2.

This grouping was rather prominent during the immediate recall. Here *B* invariably resorted to vocal-motor-auditory imagery first, and if one syllable failed to suggest the next he became conscious of the *Aufgabe* in terms of muscular strain, or of tactual-motor imagery of the table, the cards or of previous syllables. Tactual-motor imagery of syllables functioned only when auditory-vocal-motor imagery was obscure or lacking.

In the delayed recall the syllables appeared both in auditory-vocal-motor and in tactual-motor fashion. Here the latter processes were somewhat more pronounced than in the immediate recall and functioned both to supplement incomplete auditory-vocal-motor imagery and to correct wrong auditory-vocal-motor processes. The cues which aroused the imagery of the various syllables were not always obvious. The *Aufgabe* to recall readily sufficed to suggest the first syllables of the series. Usually the next three or four syllables came readily in auditory-vocal-motor fashion. Toward the latter part of the recall the syllables appeared less easily and there were frequent returns to a consciousness of the task. As far as one syllable was recalled by means of observable cues these cues consisted wholly of auditory, tactual and kinæsthetic processes.

B learned the syllables, on the average, in six repetitions of the series.

Summary of Results from Observer *A*

A's procedure in the learning was as follows: He first inspected each syllable in tactual-motor fashion but paid no attention either to the tactual or to the motor elements as such. As soon as the points of each letter were felt, there appeared visual imagery at his finger tips; this imagery became colored as the points of the letters became grouped into characteristic formations. In other words *A* invariably perceived and recognized the letters in terms of colors which took on various shapes and sizes according to the letter-meanings. Secondly, he constructed auditory-verbal imagery of each syllable, together with a color for the syllable as a whole. At times the initial letter of the syllable determined the color, and at times each letter was represented. In other instances the color of the syllable was dominated by the vowel sound. Thirdly, he retained these syllables by fixing these colors in certain positions in a visualized schema. This process was involuntary at the outset but at once became the tool of a voluntary effort to remember the syllables. Fourth, he used this schema in anticipating the syllables in the later presentations of any series. Meanwhile tactual-motor perceptions and auditory-vocal-motor imagery were used to fill gaps in this form, to correct erroneous colors or to eliminate confusions in the colors.

The immediate recall consisted chiefly of translating the colors of the form into vocal-motor imagery. Where a syllable was represented by only one color *A* was forced to resort to vocal-motor expedients and to a trial-and-error method of completing the syllable. Where the appropriate color failed to appear and to assume its correct position in the form, whether this color stood for an initial letter or a vowel, or whether it represented an entire syllable in terms of its initial letter, *A* was forced to try such indirect methods or to 'wait' until the appropriate colors appeared in consciousness. A color which occupied sufficient space in the form to cover the positions for three letters stood for an entire syllable. On the other hand certain syllables were represented by three colors in which case the recall was easy.

Colors which were not present in the form at the outset of the recall invariably appeared in the lower half of *A's* field of vision, below the form, and migrated to their appropriate places in the schema itself. Auditory-verbal imagery always brightened the colors and made them more persistent.

The delayed recall resembled the immediate recall in most respects. The colors of the form were as a rule quite as vivid as on the preceding day. Here, however, the process of translating some of the syllables wás more difficult owing to the fact that the associations between a single color and the remaining letters of the syllable had been lost temporarily. For the time being no amount of effort resulted in obtaining further cues. In such instances it invariably turned out that the color was not faithful to the syllable or letter, that a certain combination of letters in a particular word caused *A* to visualize a foreign or unrecognizable mixture of colors, or perhaps a certain color had faded or had otherwise changed in such fashion as would lead it to suggest either more than one letter or the wrong letter.

A required from 6–9 presentations in learning the syllables. *B* learned them more rapidly. This fact is explained on several grounds. (1) In *A's* case the letters thus grouped into syllables in meaningless fashion tended to produce detached or dissociated colors which did not 'hang together.' As a result his consciousness at times became flooded with 'floating' colors whose combinations were unstable. This difficulty arose even during the process of learning itself and was the source of not a little confusion. (2) The colors for syllables were often dominated by the color for one letter, thus making it necessary to resort to such expedients as tactual-motor or auditory-vocal-motor imagery in retaining the other letters; but as fast as *A* could fill out the remainder of the syllable in this fashion, the dominating color would persist in blotting out the less stable colors of the remaining letters. (3) In the act of anticipating the syllables many errors were made which had to be corrected. These errors were due to tendencies to anticipate the brighter or more stable colors out of their order in the series. (4) Attempts to fill in gaps in his form or to anticipate syllables in auditory-vocal-motor imagery inevitably dragged into consciousness additional arrays of colors which either had to be differentiated from the correct colors, in case of auditory-vocal-motor guesses, or which had to be ignored. Thus *A* was prevented from employing auditory-vocal-motor imagery in the learning unless he were certain that this imagery was correct. This difficulty was due, in general, to the fact that *A* cannot think letters, verbal fashion, in the absence of colors.

Observers *A* and *B* Compared

1. Tactual and motor images other than auditory-vocal-motor were employed by *B* in the process of perceiving letters, in retaining letters whose sounds had not become familiar, and in retaining syllables whose sounds were difficult to pronounce. This same imagery was used under similar circumstances in anticipating and in recalling the syllables. In *A's* case such imagery functioned only as a stepping stone in the arousal of colors and in no instance was it attended to as such.

2. Kinæsthetic imagery and incipient movement were employed by *B* in grouping the syllables into a schema. This grouping was done by *A* in terms of a visual schema in which each syllable assumed a definite place and took on a definite coloration.

3. *B* used auditory-vocal-motor imagery in retaining, in anticipating and in recalling the syllables. This type of imagery dominated throughout *B's* procedure. *A* used auditory-vocal-motor imagery only in connection with visual synæsthetic colors, either in an effort to arouse colors, to strengthen them, or to translate them into language.

4. Visual imagery appeared only rarely in *B's* case and was never used in the learning. In *A's* case visual imagery assumed a wide variety of rôles: (1) As synæsthetic features of his 'tactual-motor' perceptions of the letters, in which cases the colors identified the letters; (2) as synæsthetic features of vocal-motor-auditory imagery of syllables or letters; (3) as detached symbols for letters or for syllables in the absence of other processes; (4) as sections in a schema or form, standing for a syllable or an entire series of syllables; (5) as cues by means of which syllables or letters were anticipated; (6) as cues by means of which syllables or letters were retained and recalled.

Summary of Results from Auditory Presentations

B's procedure in learning nonsense syllables with an auditory presentation was confined wholly to the use of auditory-vocal-motor imagery. Upon perceiving the syllables he translated them into auditory-vocal-motor imagery and repeated them over and over again, holding as many syllables in consciousness as possible before the next syllable was given. This method was soon changed to that of anticipating the syllables, auditory-vocal-motor fashion. Whenever one syllable failed to suggest another he found himself recalling syllables which came earlier in the series or anticipating in advance such syllables as came later in the series. If these expedients failed he became absorbed momentarily in masses of motor imagery and strains which constituted effort and impatience.

A's synæsthetic phenomena began to function at the outset in learning syllables by the auditory method. Auditory perceptions consisted in large measure of flashes of brightness and of color, determined either by the initial letter of a syllable or by some other prominently sounding letter. These flashes of color ultimately became fixed in a form as in the case of tactual presentation. *A's* visual imagery became fixed by repetition as did *B's* auditory-vocal-motor imagery.

As before, those syllables were learned first whose colors were brightest and the symbolic representation of such syllables appeared as shoulders or peaks in the schema. Anticipation of syllables and the recalls took place as in the previous experiment. When this observer was about to recall any given set of syllables he was utterly lost unless the colors appeared in connection with the intent to recall. In such instances in which the colors were stable *A* found it quite as easy to repeat the syllables backwards as it was to repeat them in the order learned. On the other hand, his method of employing colors caused him numerous difficulties.

We have already mentioned several reasons for this. In the present experiment other difficulties came to light. (1) The color for the initial letter of a nonsense syllable may not be as definite, as well saturated or exactly the same quality as the color which ordinarily stands for that letter for the reason that the letter itself represents an entire syllable. Certain letters are known not only by means of their hue or brightness but by means of their shape. In case a letter represents an entire syllable its shape is changed to cover, in the schema, sufficient space to be occupied by three letters. (2) The colors for letters are influenced by the setting in which those letters appear, and by the amount of emphasis which the letter happens to receive in any particular context. (3) Certain letters take on varying colors or shades of brightness according to the different situations in which they function.

On the whole, however, the colors employed by *A* in these investigations agree with the colors for these letters, reported two and three years previous. But any variations in colors so far discovered in our investigations clearly point to the fact that where meanings remain the same so do the colors and where meanings change, so do the colors.

In the absence of a color or definite brightness, *A's* implicit *Aufgabe* to recall a given syllable consists of visual fixation upon a vacant section in the form. First his line of regard is rigidly centered upon the extreme left-hand portion or edge of this vacant section and is then moved across toward the right.

This act may be repeated several times. Meanwhile the entire form tends to split up, requiring considerable eye-movement and vocalization to keep it intact and in the center of the field of vision. Any fluctuation in shape or incipient change in color or brightness at once claims *A's* attention. Perhaps this incipient change is the forerunner of a definite color or brightness which leads to the desired verbal process and hence to the desired recall. Perhaps the vacant section in the form fails to 'fill in.' In either case the change comes of itself. Voluntary effort is invariably an indirect procedure consisting of motor-visual factors such as eye-strain, eye-movement, sudden motor-visual fixations upon changes of color and brightness in the form itself, strains in the throat, chest and arms.

On the whole, then, *A's* procedure in this experiment did not differ materially from his procedure in the tactual series other than in the method employed in arousing the original colors. In either case learning took place in terms of the manipulation of colors and these colors were determined by the letters and not by the mode of presentation. In both instances the colors were identical with those which symbolize the same letters in *A's* everyday mental life, with the exceptions above mentioned.

B's method in this experiment was likewise similar to his method in the previous experiment with the exception that, in the latter, tactual-motor processes figured in retaining, in anticipating and at times in recalling such syllables as failed to appear in auditory-vocal-motor fashion.

A comparison of *A's* performances with *B's* in this latter experiment confirms our results from the tactual presentations. In *B's* learning, one auditory-vocal-motor process led to another. These processes were his only cues. *A* had no tactual imagery. *A's* auditory-vocal-motor imagery was used only indirectly. In *A's* case colors were used at times as cues in arousing auditory-vocal-motor imagery and the latter was used as a frequent cue for the arousal of colors and for the strengthening of colors. There was no analogous process in *B's* procedure unless it could be said that one auditory-vocal-motor image suggested the next and that the next in turn reinforced the first. At any rate, in *A's* case this mutual reinforcement or facilitation process which went on between colors and inner speech is a characteristic feature of his synæsthetic phenomena and a function which strikingly reminds us of reciprocal innervation.

In this facilitation process we find another characteristic of synæsthetic phenomena. Both the auditory-vocal-motor process and the visual image play the double but successive rôle of stimulus and response. For example, the appearance of a brown color in *A's* visual form sets off the vocal-motor image of 'm' whereupon the brown color itself becomes more stable and definite. On the other hand if a certain color fails to develop to that stage of definiteness which will enable the reagent to translate it into its appropriate vocal-motor associate he resorts to vocal-motor imagery of various letters in order to find the associate which will definitize the color for future use. At times *A* finds himself vocalizing in order to arouse a desired color. Here the auditory-vocal-motor process seems to possess a stimulus function and the color becomes the response. The transition from visual to vocal-motor image is the slower of the two and is a process of translating the visual symbol, now detached from its vocal-motor associate, into the vocal-motor process. The transition from vocal-motor to visual process is exceedingly rapid and consists only of a shift of attention from the vocal-motor to the visual feature. The former seems to 'drag in' the latter. The vocal-motor feature, however, is never definite in itself and is known to the reagent only in terms of the characteristic behavior of the visual associate. By this latter he can tell definitely when the former appears and disappears. The vocal-motor-auditory feature does not exist alone in consciousness; it is noticeable whenever the visual feature is present to interpret it. On the other hand the visual feature often stands alone in consciousness. Such detached visual associates float about in *A's* field of vision and are controlled by eye-movement. By having recourse to the vocal-motor-visual combination the visual feature is rendered stable.

The two sorts of voluntary control differ in that arousing an auditory-vocal-motor process by way of the visual begins with a tentatively dissociated visual image while the act of recalling a visual image takes place at once with the recall of an auditory-vocal-motor process. The latter is never dissociated from the visual accompaniment, *i.e.*, it cannot stand alone in consciousness. The direction of *A's* attention as determined by the implicit purpose of the moment directs the order of appearance of verbal and visual process.

Summary and Conclusions

1. Synæsthetic phenomena appear throughout the learning, in *A's* case, as follows:

(*a*) in perceiving the letters and syllables;

(*b*) in retaining letters and syllables during the presentation of the series;

(*c*) in anticipating letters and syllables during the presentations;

(*d*) in immediate recall;

(*e*) in delayed recall.

2. Synæsthetic imagery functions in part as the content of self-imposed tasks. These tasks may be conscious or deliberate or they may be implied in the behavior of the reagent.

3. Synæsthetic visual imagery delays the process of learning owing to the fact (*a*) that a syllable is frequently represented only by the color of a dominant letter, (*b*) that auditory-vocal-motor guesses cannot be made without increasing the variability and complexity of this synæsthetic imagery, (*c*) that the visual associates become detached from their parent processes and fail to arouse the latter during recall, (*d*) that synæsthetic colorations vary in different situations and contexts.

4. *A's* learning has been checked in every respect against *B's* and throughout no functional differences are apparent. A comparison of the two observers shows that such differences

as are found between them are concerned only with contents.

B perceives letters tactual-motor fashion. Here the spacial grouping of letters develops in terms of bars and angles with the aid of kinæsthetic imagery and incipient movement. *A* perceives letters with the aid of synæsthetic visual imagery. Here tactual and motor factors always remain in an obscure and indefinable background.

B perceives syllables in terms of verbal imagery. *A* perceives them in terms of verbal processes plus synæsthetic visual imagery. Here, as well, the auditory-vocal-motor features remain in the background.

B perceives syllables directly in auditory fashion. *A* perceives the spoken syllables in terms of flashes of color and brightness.

The learning progressed in like fashion in both reagents as far as its functional aspects are concerned. Both were doing the same sorts of things throughout but where *B's* mental contents were tactual, motor, and auditory, *A's* were dominantly visual. The fact that *A* uses stereotyped visual imagery and the fact that he is synæsthetic resulted in certain difficulties of learning on the one hand and in a certain ease of recall on the other which were not characteristic of *B*; but these features of the learning turned out to be incidental rather than inherent in the learning process as such. They were due to the differences in mental contents between the two reagents and not to differences in basic mental functions. Such functions as attention, voluntary control, retention, anticipation, recall, association, the use of schema and the like took place in identical manner in the two reagents, barring incidental differences found in the behavior of differing mental contents.

5. *A's* synæsthetic phenomena are here identical in perceiving and in remembering, barring such changes as occur with changes in meaning.

6. We conclude that, in the process of learning, *A's* synæsthetic phenomena are normal functions whose contents,

alone, are unusual. These phenomena are as essential to his learning as processes of perceiving and recalling are essential to *B's* learning. Indeed, *A's* synæsthetic phenomena are his methods by which he perceives and recalls.

7. Synæsthetic phenomena are not confined to the realm of perception alone in *A's* case. They are quite as evident in his processes of thinking. They are throughout concerned with the development of meaning. In this case of learning meanings were confined to letters and to sounds.

8. Two sets of factors in the behavior of *A's* synæsthetic phenomena point to the view that synæsthesis is a form of conditioned reflex: (1) The parent tactual, motor or auditory-vocal-motor process acts upon the visual associate in a manner functionally resembling facilitation; (2) the colored associate is a stereotyped response to the parent process as a stimulus. The fact that the colored associate may come to function as a stimulus for the arousal of the parent process does not seem to militate against this view since the conditioned reflex as a type phenomenon need not be confined to the realm of external stimulus and explicit muscular response.

9. Voluntary control of a mental process in each reagent was an indirect process, *i.e.*, a given mental content was controlled indirectly by means of an antecedent sensory or ideational cue together with eye-movement or other incipient motor process.

LOCAL SIGNATURE AND SENSATIONAL EXTENSITY

BY W. C. RUEDIGER

George Washington University

The various theories aiming to account for local signature may apparently all be classified under three heads. These are (1) the kinesthetic theory, (2) the sensory-complex theory, and (3) the sensory-element theory.

The kinesthetic theory is explained as follows by William McDougal:

> The process by which the axis of the eye is turned toward the optical image is probably a pure reflex of the spinal level. . . . When such a movement of the eye is made, a certain complex of kinesthetic sensations is excited. . . . Hence we may assume that the turning of the eyeball in its socket to any position excites the sense-organs of the socket and surrounding parts in such a way as to arouse a certain complex of kinesthetic sensations that is peculiar to each position, and that this complex of kinesthetic sensations is the primary form of the local sign. (Primer, p. 97.)

Against this kinesthetic basis of retinal local signature Charles S. Myers urges the following weighty objections:

> In cases of congenital blindness, the individual is at once after operation able to recognize differences in size and in form of images received by the retina. So instantly is a round object seen to differ from a triangular one, that the judgment is clearly independent of the training which would be needed before the eyes could be moved in a fashion orderly enough to follow the outlines of the object.
>
> Again, in certain cases of retinitis, where the elements in the affected region are abnormally crowded together or separated by inflammatory products, a local 'metamorphopsia,' or distortion of vision, is produced. It arises from the fact that the displaced retinal elements retain their normal local sign under these conditions. A line whose image falls on the affected area of the retina, appears bent or bowed out, according as the elements are unduly separated or contracted. Such distortion would surely be corrected in time if motor sensations were the basis of local signature. In point of fact, it is said to persist for months or even for years—indeed, as long, perhaps, as the pathological condition remains.
>
> Lastly, attention may be called to the fact that lines which subtend less than an angle of one minute can be spacially distinguished by the eye. If such discrimination depended for its origin on kinesthetic sensibility, then the delicacy of kinesthesis in the orbits must be almost incredibly greater than that in any other part of the body. (Experimental Psychology, p. 230.)

On the basis of these objections, Myers thinks "we can hardly avoid the conclusion that the local signature of the retina is innate, and that the basis of it is to be sought rather on the sensory than on the motor side of the retinal sensori-motor mechanism." A similar conclusion is probably true as regards the skin.

According to the sensory-complex theory the sensations received from the elemental end organs in the skin, from the end organs of the kinesthetic and visceral sensations, and from the separate rods and cones in the retina are all respectively alike, and the differences in local signs that we experience are all due to different combinations or complexes of these homogeneous elemental sensations.

According to the sensory-element theory, the elemental sensations received in the sense departments just mentioned are all slightly different, and specific locations on the body, within the body, and directions in space perceived visually, are learned by means of associations with these differences. The theory assumes that no two elemental skin, retinal, or organic sensations are really alike, and that each possesses the earmarks of the particular region of the body from which it comes. The difference is not caused merely by differences in the structures accessory to the receiving nerve ending, such as the texture of the skin and the proximity of bones, but is one that inheres in the sensation itself. This may, perhaps, be best understood by assuming that the sensory areas in the brain are made up of a point for point projection of the extended receiving areas of the skin, retinas, and internal structures. To have all touch sensations, for example, strictly alike in the touch quality, they should all theoretically be received by the same brain elements, and if the same brain elements could or did receive the touch sensations coming from all parts of the body, there is logically no reason why this receiving should not be done by one neurone. But with an extended tactual area in the brain it may easily be true that each tactual end organ has its own receiving point in the brain, and because of having its own brain localization it may also well have its own psychical characteristics so that no two touch sensations are really alike, while they are nevertheless enough alike to be called tactual. What holds for touch sensations may hold also for visual and organic sensations. The point involved may be aided by an analogy. No two people are really alike—we can tell them all apart—yet they are sufficiently alike to be classed as people. Each person has his own 'local sign,' so to speak.

Granted that the arguments advanced by Myers are conclusive against the kinesthetic theory, the choice then lies between the other two. This choice must in the end be made on the basis of experimental and anatomical evidence.

The experiments that I am reporting in this paper all appear to point to the truth of the sensory-element theory. I performed experiments only upon the skin, stimulating the radial side of the forearm in four subjects with a Bloch instrument. The subject was touched with his eyes closed and was then asked to locate the touch with the point of a pencil, eyes open. The amount of error was measured in mm.

For readiness of comprehension the experiment may be divided into four series. In one series the stimulus was applied to a conspicuous vein and in another to a portion of the skin where no vein was in evidence. Furthermore, the stimuli were given in two degrees of intensity, namely 1 gr. and 10 gr. Twenty stimuli were used in each series, making a total of 80 for each subject. The results are summarized in the following table:

Vein				Arm	
Subject	Pressure	Av. error	M. V.	Av. error	M. V.
A...........	1 gr.	8.5	4.7	14.5	6.5
	10 gr.	8.5	4.7	10.3	5.2
B...........	1 gr.	6.2	3.3	7.3	3.4
	10 gr.	6.2	2.8	6.8	3.5
C...........	1 gr.	9.3	4.2	12.7	6.1
	10 gr.	8.6	5.5	10.0	4.8
D..........	1 gr.	11.0	7.1	10.0	6.6
	10 gr.	8.0	4.6	17.0	8.3

If the sensory-complex theory were true it would seem reasonable to expect that a strong stimulus, within certain limits, would enable more accurate localizations than a weak stimulus, but this is not the case. The advantage appears to be slightly in favor of the weak stimulus.

Again it would seem reasonable to expect that when the stimulus is applied to a vein, where the subcutaneous tissue is uniform, localization would be less accurate than when applied to other parts of the skin where the subcutaneous tissue is variable. But the reverse appears to be the case. In seven out of the eight cases localization was more accurate on the vein.

In a preliminary way I have experimented also with cold spots. I spent about three hours letting one of my students stimulate marked cold spots in my forearm. It proved somewhat difficult to get cold sensations unaccompanied by touch sensations, but many pure cold sensations were received. Localization was just as accurate as with touch sensations or as with cold and touch sensations combined.

All these results, it seems to me, point toward the truth of the sensation-element theory and away from the sensation-complex theory.

There is another argument against the complex theory. A complex is possible only on the assumption that the elements that go to make up the complex are different. With like elements, only a difference in intensity could be produced and that would be of no service in aiding localization. In the skin and the tissues underneath there may be a sufficient variety of sensations to give complexes, even if all cutaneous touch sensations were alike, but it is hard to see how this could be true in the retina. But if the complex does count in the skin, this should have come out in our experiments. The fact that localization is no less accurate with a weak stimulus than with a strong one, and the fact that it is more accurate on than off a vein count against the complex theory.

Granted inherent differences in sensations from different parts of the body and we have a clear basis for local signature;

indeed, such differences are local signs. Yet such differences of themselves do not give spacial position or direction; they are merely the ground on which such position or direction may be learned. The acquisition of space relations on the skin, for example, consists of learning the meaning of these minute differences in terms of visual and kinesthetic sensations. In acquiring the ideas of visual space it is necessary for the kinesthetic and skin sensations to become associated with the retinal signs.

Sensational extensity, as an elemental quality, becomes an entirely superfluous assumption. The awareness of differences in size on the skin or in the visual field are adequately explained as perceptual complexes. The end of a pencil stimulates many more touch end organs than the point and so arouses a different sensory complex which experience has taught us to mean a larger area.

The superfluous nature of the quality of extensity in sensation does not, of course, prove its non-existence. The arguments, however, that are advanced in support of extensity as an elemental quality are not sound. Titchener, for example, says:

> Sensations of color are spread out really into length and breadth; they appear as spacial extents. And this attribute of extent is part of their very constitution. Reduce the color to a pin point, and it still occupies space; think away the spacial attribute, and the sensation has disappeared with it. So with pressures; set the point of a stiff horse-hair lightly down upon the skin, and the sensation is extended, diffused over a mental area. Certain sensations, then, have extent; others, as odors and tones, show no trace of it. (Textbook, p. 54.)

Looked at from the objective, or logical, side, this argument appears unanswerable, but when it is looked at from the subjective, or psychological side, its unsoundness becomes clearly evident. Without the power of movement and without any other experiences whatsoever, the stimulation of one end organ either in the skin or in the retina would appear to the subject merely as a modification of consciousness, and there is no reason why it should have any more suggestion of extent than a sensation of smell or sound. Extent as we know it is an objective datum that acquires its objective meaning through experience. Just because we cannot think of the external world without it does not make it an elemental quality of sensation. Any sensation coming from a single sense organ may still be absolutely innocent of it.

Another argument for the sensational nature of extensity that is frequently given is still more obviously fallacious. It is said that the fact that the congenitally blind can at once distinguish different shapes, such as circles and squares, when their sight has been restored by a surgical operation, proves that visual sensations have the innate quality of extent. These people seem to forget that a person when looking at a ball or a box has not only one, but many end organs in the retina stimulated, so that he bases his judgment, not on one, but on a complex of sensations. This complex *has* to be different for a circle from what it is for a square, and it is not at all surprising that a person even on the first trial can tell one from the other. There is no more reason why he should not be able to distinguish them than why he should not be able to tell one complex taste from another complex taste, or one complex sound from another complex sound. It is right here that the innate differences of the various single sensations, granted their existence, do their service. These should give a characteristic consciousness for each group or complex of sensations, which, when received from the skin or retina, is interpreted by the assistance of other experiences to mean certain definite spacial relations.

So far as the functioning of the mind is concerned, it makes no difference whether sensations have extensity or not, or whether some sensations have this quality and others do not. If they have it, it is an interesting fact, but that is all. Space perception, both visual and tactual, may develop perfectly well without it. Professor James is very insistent in the beginning of his chapter on space perception that all sensations have the inherent quality of extensity or voluminousness, but when later on in the chapter he explains how our actual and precise ideas of space are obtained he makes no use of this quality. Neither does any other writer make use of it.