Volume 5 March, 1941 Number 1

BF
1001
J86

The Journal of Parapsychology

GARDNER MURPHY and BERNARD F. RIESS, *Editors*
EMMETT TAYLOR, *Managing Editor*

BOARD OF REVIEW
B. F. Riess, Chairman

Lawson Dodd
B. M. Hilgard
J. L. Kennedy
Louis Long
Irving Lorge
F. H. Wilcoxon

CONTENTS

Experiments Bearing Upon the Precognition Hypothesis: III. Mechanically Selected Cards — J. F. Milne 1

Conditions Affecting ESP Performance — Burke J. Smith and Edmund P. Gibson 56

Letters to the Editors 87

The Journal of Parapsychology

A QUARTERLY DEALING WITH TELEPATHY, CLAIRVOYANCE, AND OTHER PARAPSYCHOLOGICAL PROBLEMS

The Journal is published quarterly. Contributions submitted for publication and all editorial communications should be addressed to Gardner Murphy, College of the City of New York, or to Bernard F. Riess, Hunter College, New York City. Correspondence with the editors is advised before submitting articles other than reports of experimentation. Since the Journal does not forward manuscripts by first-class mail, it cannot guarantee that they will not be lost in transit, and contributors are urged to keep copies of their papers. All contributions should be typewritten, double spaced. References should be given in the form adopted by this Journal.

The Journal does not supply reprints; authors are given two proofs in the issue in which contributions appear. Correspondence concerning subscriptions, change of address, back numbers, and other business communications should be addressed to Bernard F. Riess, Hunter College, New York City.

The subscription price is $3.00 a year; single current numbers of the semiannual issues are $1.00, of the quarterly issues, 75 cents. The rate for back volumes, when available, is $4.00; for single numbers of volumes 1 and 2, $1.00; for those subsequent to volume 2, $2.00. Missing numbers will be supplied free when lost in the mails if notice is given within one month of the date of issue. All remittances should be made payable to Bernard F. Riess, Hunter College, New York City.

The Board of Review of the Journal currently includes the following: Drs. Saul B. Sells, Louis Long, Irving Lorge, Lillian Dick, [illegible], J. L. Kennedy, and [illegible].

The editors have asked Drs. H. Hotelling and J. A. Greenwood to serve as mathematical consultants in specially difficult quantitative problems, where mathematics rather than experimental psychology is involved.

It is through the generosity of three anonymous contributors that this Journal will be enabled to appear in 1941 as a quarterly.

General
Stuchert

EXPERIMENTS BEARING UPON THE PRECOGNITION HYPOTHESIS[1]

III. Mechanically Selected Cards

J. B. Rhine
Parapsychology Laboratory, Department of Psychology,
Duke University

ABSTRACT: In an earlier number of this *Journal*, results were reported which led the author to the conclusion that subjects could, to some extent, predict card order in decks that were later shuffled by the experimenter. This was not concluded to be evidence of precognitive ESP since, as it was later shown, the deck might have been shuffled with some apparent guidance from ESP ability--the ESP shuffle effect--and that this might account for the earlier results. The present report deals with tests based on mechanical selection of cards, selection in which the ESP shuffle could have played no part.

Herein are described the procedure, results, and evaluation of 1,608 runs through the 25-card deck, under conditions called Precognitive Matching. In this technique the subject matches the cards of a deck against five blank spaces in which there are to be placed, later, five key cards mechanically selected by a fixed procedure. By eight different standard measures for evaluating the results, they are extra-chance in character. All the known and relevant ways of explaining the results are then considered but are found inapplicable. While the desirability of further evidence is emphasized, no alternative at present acceptable is found to the hypothesis of precognition; that is, the ability to respond as if to cognition of events (card-orders) not yet in existence and not rationally inferable or voluntarily produceable. The effect of rewards, grouping of subjects, and certain other conditions are experimentally investigated and discussed. An appendix is added giving a summary of results of all precognition tests to June, 1939, and a supplement to this brings the summary up to October, 1940.

---

1. Mr. Edmond P. Gibson, Record Librarian of the Parapsychology Laboratory, has participated as assistant experimenter throughout the experimentation herein reported. The extent and manner of this participation is explained in the text. He has also

Introduction

In the fall of 1933, experimenters in the Parapsychology Laboratory of Duke University, having found that distance between percipients and events perceived did not interfere with extra-sensory perception, undertook to ascertain whether "distance" in time would do so. The first report of this study was issued by the present writer in 1938 and in it were summarized the total results, 113,075 trials in pre-shuffle card-calling tests in which the shuffling was done by hand, made by 49 subjects with the collaboration of a number of experimenters. These trials gave a positive total deviation of 614 hits (average score per run, 5.14) which gives a significant critical ratio (4.5).

The procedure followed throughout the series under discussion consisted essentially in having the percipient call aloud or write down a run of 25 ESP symbols, intending them as predictions of the order which a specified pack of cards of equal number would have after it had later on been shuffled or cut in a specified manner by the experimenter.

The alternative hypotheses to ESP that were being actively considered at that time were regarded as met by the conditions: (1) The results were statistically significant and the chance hypothesis was excluded; (2) sensory cues could not have aided the subject in making his calls since the order of the cards did not exist at the time; (3) the selection hypothesis was excluded by the completeness of the reporting, which included all trials and scores made under the conditions described.

There was, however, another hypothesis alternative to precognition raised in this Laboratory as a result of some related experiments. This was called the

---

(Footnote continued) prepared most of the tables, with the independent checking either or Miss Esther May Bond or myself. In addition to these contributions, I am indebted to Mr. Gibson and Miss Bond for conducting and rechecking all the analyses and computations entering into the report. This large share in the experiments and in the report is gratefully acknowledged.

To Dr. J. A. Greenwood of the Department of Mathematics, to Dr. J. G. Pratt and Mr. C. E. Stuart of the Parapsychology Laboratory, and to my wife, Dr. L. E. Rhine, I am indebted for careful readings of the manuscript in the course of preparation and for helpful suggestions made.

General Library
Psychology
Direct
10-24-46
56149
v.5,no.1

ESP shuffle hypothesis. It implied that the investigator shuffling the pack of cards may conceivably have been guided to some extent in the shuffling and cutting by ESP awareness of the order of the cards and of the subject's record. Since the score average was low (5.14), the placing of an occasional card by some other method than chance would suffice to provide the results obtained in the pre-shuffle card calling work.

The ESP shuffle hypothesis had itself been under investigation here for a number of years when the first report on the subject was made by Smith, Woodruff and myself in the June, 1938, number of the *Journal of Parapsychology*. At that time, two series of investigations totalling 211,520 single trials were reported, giving a deviation beyond chance of 1,959 (average score per run, 5.23) which has a critical ratio of 10.4. In this general investigation of the ESP shuffle hypothesis, the procedure involved mainly the shuffling or cutting of the pack of cards with a view to making it match, in a better than chance degree, the order of another pack, (or a record of 25 symbols) designated as the target. In 68,200 trials of the total, the deck to be matched was completely screened from the subject's vision. The critical ratio of the deviation of 494 hits (average score 5.18) in this more safeguarded block of data alone was 4.6.

So far as the averages were concerned, it was clear from the work on the ESP shuffle hypothesis that results could be obtained by that method which were as good as and even better than those reported for pre-shuffle card calling. The conditions of the latter experiment were in every respect equal to those of the earlier report. There was still no evidence that the ESP shuffle hypothesis was *necessarily* the explanation of the pre-shuffle-card-calling test results, but it appears to be a clear possibility. Further investigation was evidently required before the precognition hypothesis could be brought to crucial examination. The step required would have to involve the elimination of the possibility of arranging the cards by means of the ESP shuffle to meet the predictions of the subject.

Mechanical Selection Background

The required step was taken in 1934, when attempts were made to transfer the shuffling of the cards entirely to mechanical devices and thus eliminate the possible personal factor of the ESP shuffle.

Considerable difficulty was encountered in the development of suitable mechanical shufflers and in finding suitable card materials adapted to the shuffler. Finally, however, this was overcome to the degree that systematic investigation of the precognition hypothesis with the aid of mechanical selection of the card order of the targets was undertaken. In a large portion of this work, the earlier procedure was followed with the exception of the manner of shuffling; that is, it was still pre-shuffle card calling. The designated deck would then be put in the shuffling machine, given the prescribed number of revolutions, removed, recorded and checked for correspondences with the subject's predictions.

In other lengthy series, the procedure designated as Precognitive Matching was followed. A deck of cards was matched by the subject against five blank spaces on the table in which later five selected *key cards* would be placed. The subject attempted to place each card opposite the space that would later contain that one of the five key cards to which he thought it would be similar. The key cards were taken in a prescribed manner from a mechanically shuffled deck, the shuffling having been begun after the subject had laid the last card on the table. In placing the 25 cards of the pack opposite the 5 spaces intended for the key cards, the subject was in effect attempting to predict what the key cards would be. The random factor was introduced by the shuffling machine and if extra-chance coincidences resulted, the precognition hypothesis (that reliable responses may be made to a future random order of events) would presumably be supported. (See below under "General Conditions and Procedures" for further details of this technique.)

In place of the shuffling machine which directly shuffled the cards, other methods of introducing the random element were used to a considerable extent. For example, a number of dice were thrown from a cup to determine the point at which the target deck of cards was to be cut. In another method the totals of the faces of dice similarly thrown were taken to indicate one of a number of envelopes, each containing a set of key cards to be used in precognitive matching tests. The particular envelope selected was one from a large number containing randomly distributed keys; the throw of the dice made the selection of the particular envelope itself a random one.

In another series the dice were thrown mechanically, and in still others such devices were used for cutting target decks of cards (against which predictions of the subject were directed) as the use of figures to appear in the newspaper on a day specified in advance--e.g., those given by the maximum and minimum temperature, or the last five digits on the right for the total stock sales for a given market.

During the intervening years since the first attack was made on the precognition hypothesis with a view to escaping the application of the ESP shuffle hypothesis, these methods and minor modifications of them have all been used in one series or another of ESP experiments. But as those who have followed the developments of the ESP research field are well aware, advances in precautions have continued to be made from time to time in the ESP test procedures and as a consequence much of the work done with the procedures just mentioned does not appear today to meet explicitly all of the most recently advanced possible counter hypotheses. It does not, for example, equal in safeguarding conditions the most recent work represented by the report from this Laboratory of Pratt and Woodruff in a recent number of this Journal. This is not to say that it does not represent evidence on the question of precognition, and evidence that must be taken seriously; probably all of it will be found to have made some contribution to the answering of the question, but in view of the enormous amount of work involved in the re-checking, cross-checking, and making of various analyses that would be expected in a full utilization of this material, it does not seem at this time a wise expenditure of energy to attempt a full report--not, at least, until more recent evidence obtained under the most up-to-date conditions has determined whether the question as a whole warrants the extended consideration which this evidence might afford. (There is, however, provided below in Table VI, under the discussion of the question of "Selection of Data," a brief summary of all the data acquired up to the end of the year in which the work of this report was done.)

In the present report, only that work based on mechanical selection of cards which has been conducted under the most advanced experimental conditions will be included; and of this, only that which was conducted by Mr. Gibson and myself. The work of comparable character conducted by others is in process of analysis and preparation for later reporting.* The general question

*Miss Hutchinson's report, as it turns out, preceded this one (see the December, 1940, issue) but was written after it; hence the omission of other reference to it.

of the background of experimental results from which the material for this report is taken will be discussed fully below.

Introductory Series

The research program covered in this article had its initiation in two exploratory series. These series are of sufficiently different character from the principal work of this report that a separate (though brief) account of these introductory series is necessary. The grounds for regarding these separately will be clarified in the discussion of the question of selection of data.

Introductory Series A. This series contained 80 runs (of the total of 160 runs made) that were of the precognitive matching type, in which the key cards were selected by means of six *hand thrown* dice. The upturned faces of the dice were totalled and the number so found was used to indicate which of a number of envelopes containing random distributions of key cards was to be used for a given run. The point of interest in this work lies in the fact that although the non-precognitive part of the series gave a positive deviation, the precognitive section was a *negative* deviation. This negative deviation was small, -12 for the entire 80 runs (average score, 4.85), and was of course insignificant.

But, small though the deviation was, it seemed worthwhile to continue the experiment with a view to ascertaining if there were associated with the particular type of test (i.e., precognition) some score-depressant factor. (This interest was furthered considerably by the negative deviation which I had just reported in connection with the enclosed card tests in which there had been delayed checking (7). It was realized, too, that a consistent negative deviation would serve quite as well as a positive one as a means of exploring the hypothesis of precognition and might be more practically and repeatedly obtainable.

Introductory Series B. The next experiment was accordingly planned to consist entirely of precognition tests with the hope of accentuating the hypothetical score-depressant factor. Four hundred runs were planned in all, and the procedure was that of the precognition tests in the preceding series (A). The principal fact is that the series as a whole likewise gave a negative deviation of -23 (Average score, 4.94).

It seemed warranted again that this experimentation be continued under essentially the same general conditions. There had been introduced, during the series, the offer of small rewards, such as movie tickets, for the achievement of certain scores. This practice made the tests much more interesting to the subjects and was adopted as one of the conditions for the projected series.

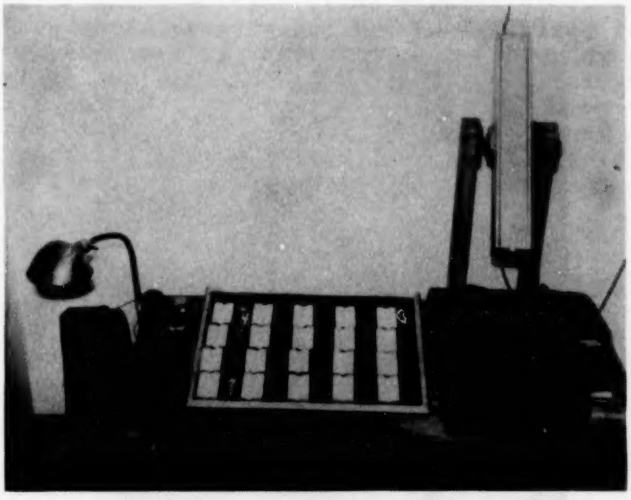


FIGURE 1a. THE APPARATUS USED FOR THE SELECTION OF KEY CARDS

On the extreme left is the rotating dice machine, driven by motor. Next is the large tray with four small trays, one for each set of key cards. At the right is the motor-driven shuffling machine. The cards are placed in one end of the long box at the top which, in its rotation, throws the cards over wavy surfaces within and thus shuffles them. On the extreme right is the box of opaque envelopes containing sets of key cards for selection by the dice machine.

The Mechanical Selection Experiments

The experiments with which this report is directly concerned consist of two divisions which will be called the Adult Division and the Child Division. Each of these consists of several series which will be

numbered in chronological order. The four series of the Adult Division experiments are those which grew out of the interest in the hypothetical score-depressant factor suggested by the introductory experiments described above. The three series of the Child Division were introduced as a comparative study, a control on the factor in question, and were carried along as a parallel investigation. The difference (besides age of subject) between the two divisions lay in the expectation regarding the hypothetical depressant factor.

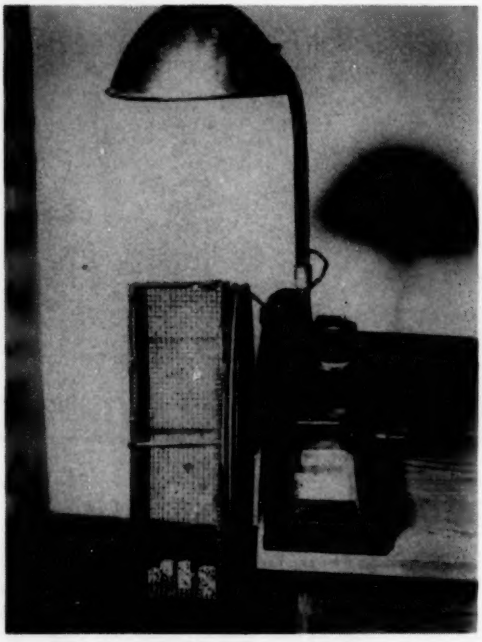


FIGURE 1b. MOTOR-DRIVEN, ROTATING DICE MACHINE.

Adult Division. For the particular group of adult subjects who had been used in the Introductory Series (and who were to be used, as far as possible, in subsequent series), there seemed to exist some ground for expectation of a negative deviation. Earlier ESP research had revealed instances of prolonged scoring below chance expectation. The student of ESP is therefore not unprepared for this effect. I had, myself, reported experiments in which subjects had deliberately produced

negative deviations (8), and other experiments in which, even with apparent effort at scoring above chance, negative deviations were nevertheless produced (7, 8). Pegram (5) confirmed the fact of voluntary control in producing positive and negative deviations at will.

Child Division. There was no such ground (from earlier experiments), however, to expect such a score-depressant factor to operate in the instance of child subjects. It had not been reported previously as occurring in tests with children and as a further guide on the question, I had the unpublished results of an incomplete series conducted by my wife, Dr. L. E. Rhine, which showed positive results [282 runs with a positive deviation of 57 (average score, 5.20)] in tests of children with precognitive matching using small rewards such as were used in Introductory Series B. Therefore it seemed advisable to run parallel series, using subjects who gave the earlier results of both types as far as possible.

General Conditions and Procedures[2]

Since identical objective conditions and procedures were to be followed both for the Adult and Child Divisions, the same description will do for both. The Adult Division consists of four series and the Child Division of three, a difference due to difference in number of subjects participating.

Experimenters. Two experimenters were present on every occasion and during two series a third observer was added for a purpose to be specified below. I, myself, was present throughout all series except for rare and momentary absences to take telephone calls, during which time my secretary, Miss Marion Barber, who is thoroughly familiar with the procedures and requirements, substituted for me. Mr. E.P. Gibson, Record Librarian of the Laboratory, was assistant experimenter in all series with the exception of an occasional run or short series of runs, during which time a member of the experimental staff of the Laboratory took his place. When a third observer was needed, either Miss Barber or Mrs. Dorothy H. Clark, general secretary to the Laboratory, served in that capacity. The responsibilities of the different experimenters were so distributed that no one individual could accidentally or with motivation affect the results in favor of any hypothesis except that of chance. The specific tasks and responsibilities of each experimenter will be described.

2. A diagram of the general plan of the experiments is given in Figure 3.

Programs filed in advance. According to the rule of the Laboratory, an outline of the research program was filed in advance of the research, one copy going to the Record Librarian and the other filed in the experimenter's record book. This advance statement set the length of each series, stated the procedure and conditions to be followed and outlined the safeguards to be emphasized.

The subjects. As far as possible, the subjects used were the same as those who had given the results in the previous series upon which these experiments were based. The 19 subjects used in the Adult Division had either been used in the earlier series (Introductory Series A and B) or were transient subjects used in one experimental session only. There was no serious effort made at selection of subjects other than to give major emphasis to the group previously tested. These subjects, as it happened, were also most convenient to obtain, since they were mainly associated with the Laboratory staff. Most of the other subjects were introduced in a session given over to demonstration of the procedure to a visiting group. Five of the child subjects used earlier by Dr. L. E. Rhine did all of the work in the Child Division. Four of these were children of staff members and were selected because of their ready availability for repeated testing.

Test materials. The commercial ESP test cards in packs of 25 were used throughout. Decks were counted from time to time--in certain series at the end of each run--making sure of the number of cards in the deck. As soon as decks showed any wearing effects from use, they were replaced with new cards. In view of the nature of the tests, the defect in the printing of these cards could afford no basis for concern and no screening was needed.

For the shuffling of the key cards, a shuffling machine of the type described by Smith (12) was used throughout the series. For the throwing of the dice used in determining the selection of the key cards (see below) a motor-driven, rotating dice cage was used. This apparatus, together with the trays used for holding the key cards and the opaque envelopes, are illustrated (see Figure 1a).

Laboratory arrangement. Two rooms of the Parapsychology Laboratory were used throughout the work here reported. The rooms are adjoining, but are separated by two brick walls which are themselves separated by an air

space. There is no connecting door. An electric buzzer connects the two rooms, enabling the experimenter working with the subjects in room 3B-b (See floor plan in Figure 2a) to signal the second experimenter (Mr. Edmond P. Gibson, hereinafter referred to as E.P.G.) when the subjects had completed a given run. The buzzer was used only for that purpose. It will be seen from the diagram that the experimental room contains a long table. It was this upon which the card matching was done. Places for four sets of key cards were indicated down the center line of the table by sets of blank face cards permanently fastened to the coarse black cloth table cover. A maximum of eight subjects could work at one time if desired. The shuffling and dice throwing machines, as well as all cards from which the selection of keys was made, were in the adjoining room, 3K-d. During the experimental sessions, the doors of room 3B-b were kept closed and there was no communication between the two rooms, with the exception of the one-way signalling mentioned above.

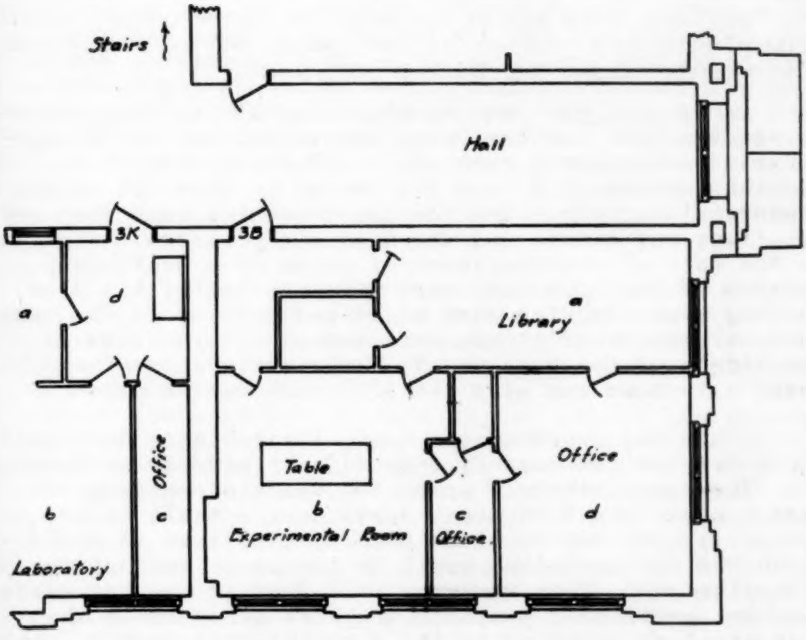


FIGURE 2a. FLOOR PLAN OF ROOM IN WHICH THE EXPERIMENTS WERE CONDUCTED.

3B-b is the test-room. All the tests were conducted on the long table indicated. 3K-d was the location of the card selection machines. They were on the table shown. The doors 3B and 3K were left open but those of 3B-b were kept closed.



FIGURE 2b. EXPERIMENTAL TABLE AT THE CHECKING STAGE.

Key cards (four sets of five each) are in the center. Eight decks of cards have been matched, two against each set of key cards. Four subjects and two experimenters are present.

Procedures. At the beginning of the experimental session both machines were turned on, the shuffling machine containing a pack of 50 ESP cards, the dice machine containing 6 dice 3/8 inches in size, of common commercial variety. The shuffling machine made four revolutions per minute and the dice machine, six. Figures on the rate of rearrangement of cards by a shuffling machine of the type used were given by Smith; the dice, rolling over the wire mesh sides and baffles of the cage, obviously received an adequate amount of disturbance. The length of the cage was 15 inches; the size of the mesh, ¼ inch on one side and 1/12 inch on the other.

In the experimental room, the subjects were told the nature of the test if they did not already understand it. They were asked to match the decks of cards given them against the five blank cards on the table before them; i.e., to try to place each card so that it would match the key card that would be placed on that space after the run. They were not told whether the key cards for that particular position would be selected by the mechanical shuffler or by the dice throwing machine, but they knew that it would be by one or the other method. After receiving an explanation of the system of reward for high scores, they were told to proceed to deal out the cards, keeping them in the inverted position.

In view of the fact that there were at the time no key cards in place and the further circumstance that the key cards were to be selected at random, no special effort was made by the experimenter to prevent the subject from obtaining cues from the cards; that is, no screens were used, no attention given to lighting effects, and no special scrutiny given to detect any attempts by the subject to turn the cards so as to see their faces. The possibility of such effects entering into the experiment were fully in mind, but it was recognized that the expected score average could not be influenced in that way. If sensory cues from the cards influenced the subject in placing them, there might, it was realized, result an altered *distribution* of scores; but if such were the case, the statistical analyses of the results would reveal the effect and distinguish it from score distributions resulting from ESP matching (see below under "Stacking Effects").

Likewise, little attention was given to elaborate shuffling of the decks of cards used by the subjects, except with the introduction of new decks. After a deck of cards had been distributed and the run finished, the deck was considerably mixed up as it was assembled from its spread-out position on the table. This mixing up was especially effective since each card lay in an isolated position at the start. After the cards had been gathered up and inverted, they were frequently turned over to the subject without a cut or further shuffle. The same considerations of safety, however, enter here as in the paragraph just preceding.

If only one subject was present, he was asked to match eight decks at a single round. The number of subjects working at a time varied from one to eight. If only two were working, effort was made to prevent them from working opposite each other at the same time because of the possible distracting effect of the responses of one upon the other. As soon as all the decks were placed or matched, the buzzer signal was given to the second experimenter, E.P.G., who at once stopped both machines. In the experimental room, the first experimenter, J.B.R., with the aid of the subjects, turned the matched cards face upward, laying them out so that they could be easily identified. Toward this end, cards of the same smybol opposite a given key space were grouped--i.e., placed side by side. This experimenter, in all but two of the seven series, then proceeded to make records of the card order on consecutively numbered record sheets issued specifically for the particular experiment and

stamped with the seal of the Laboratory. These records (made in ink) had to be completed before the opening of the door to admit the second experimenter, E.P.G., with the key cards.

Experimenter E.P.G., after stopping the machine, opened the shuffling machine, removed the deck of cards (if any cards were out of place, they were picked up as they lay and placed on top of the deck). The top card of the face down pack was taken to be the key card for space No. 1 for the first set of spaces at the west end of the table used in the tests; the next card down of different suit, for key card No. 2; the third card of different suit, for space No. 3, etc., until the five key cards were obtained for the first set of spaces. When the first set of key cards was made up, the next card in the mechanically shuffled deck was taken as the first key card for the next set, going eastward on the table. The next following card of different suit was taken for the second key card, and so on until the second set was completed. Thus two sets of key cards were selected from the deck of 50 mechanically shuffled cards.

With the exception of the first few runs made before the dice throwing machine was adapted for the purpose, the other two sets of keys (at the east end of the table) were selected as follows: The upturned faces of the six dice were totalled by the second experimenter. (Special provision was made for instances in which there was some irregularity--such as a die being on top of another. In that case, another revolution was given.) The number arrived at was used to select an opaque envelope from a box (containing 36 or more such envelopes) kept in readiness for the purpose with one end identified as the front. Beginning at the front, the experimenter would count off envelopes until he came to that which corresponded to the number obtained from the dice. This envelope contained five key cards and these were taken from the envelope in a prescribed manner; the top card, when the envelope was held so that the keys were face down, became the first key card for position No. 3 from the west end of the table, and the other cards of the envelope followed in sequence. The envelope immediately following the first one selected provided in like manner the key cards for the fourth position (set of spaces). The key cards had been selected and shuffled before placing them in the envelopes by E.P.G. and an assistant, working together to check against error. A large stock of these envelopes was prepared and kept in reserve, the box being replenished as needed. As a set of key cards was used it was reshuffled, returned to its

envelope, and this was stored with a stock of 200 such envelopes from which the needed supply was drawn. In the shuffling of the key cards for use with the dice selection method, no special criterion or set of precautions was introduced in view of the fact that not only do other random factors enter into the final production of scores but an eventual check-up can be made statistically which will determine whether any error could have been introduced from this source (see the cross-check results).

Four trays, each having five shallow boxes, were used for the key cards; each was numbered and each had its proper east-west orientation indicated in order that reversal of orientation could be detected at once if it occurred. The key cards were placed face upward in the boxes of the trays immediately upon their selection. The key cards were recorded in ink on comparable numbered and stamped record sheets. Immediately following the completion of the selection and the completion of the recording of the key cards by E.P.G., the four trays were carried on a large tray to the experimental room. On entering the experimental room (after the recording in there was completed) E.P.G. immediately put the trays with the key cards in their appropriate spaces. Experimenter J.B.R. recorded on the same record sheet on which the target decks were recorded, the appropriate set of key cards against which they were to be matched. He then checked the record immediately and determined the score for each deck. Meanwhile, experimenter E.P.G., retaining his own record of the key cards, proceeded to check the number of hits in each target deck, checking directly from the cards. He recorded his scores and checked these against the corresponding scores of the other experimenter. The subjects also invariably made an independent count. If there was any discrepancy, the cards and the records were immediately consulted. The recording techniques were definitely directed toward rendering the scoring by the two experimenters independent.

J.B.R. then transferred his scores to his personal record book to be kept in his possession until the end of the series. After that he turned over to E.P.G. his record of the key and target cards on which his scores were recorded. Thus there was not only recording of key and target cards before the two were brought into the same room, but independent counting, one from the records and one from the cards; and finally, also, there were independent records of all scores retained by the two experimenters. In the event of any later discrepancy or evidence of error found in the handling of the records,

the rule was laid down that the score nearer the chance mean should be taken. Throughout the period of checking, from the point at which the key cards were brought into the experimental room, experimenters and subjects were required to refrain from touching the cards on the table.

During two of the seven experimental series, in order to speed up the procedure, the recording of the two sets of cards--key and target cards--was eliminated and a third observer was introduced; thus a triple count of the card coincidences could be made directly from the table. The two regular experimenters each made his independent count and recorded the scores immediately in his record book, checking with each other at the end of the count. The third counter made a record of the number of hits in each column of cards opposite the key cards. These were then totalled from the five columns for each run to get the score. This was then checked with the other two experimenters. In the presentation of the results, the two series having the triple count instead of detailed recording will, of course, be indicated and full comparison of results made possible.

Special precautions. Most of the special precautions have been mentioned, but they will bear reviewing. In order to avoid the possibility of loss or exclusion of any record, consecutively numbered record sheets were used throughout and a record of these was kept by the Record Librarian at the time of their issue. In five of the seven series, independent records of the two sets of cards to be matched were made before the two sets of cards were brought together, and in the exceptions triple and independent counting was substituted. Frequent checks were made on the number of cards and the composition of the decks which were used. Independent score records were kept by the experimenters throughout the series. Comparisons of scores were made immediately following checking and if any discrepancy occurred, it was immediately cleared up. The card records were re-checked at the completion of the series and in the event of discrepancy with the score records kept by the experimenters, the score favoring the chance hypothesis was taken.

Experimental atmosphere. No one in the adult series was told, and it seems safe to conclude that no one discovered during the period of the experiments here reported, that the series of experiments was planned to test the hypothetical score-depressant factor--that is, that negative deviations were being sought. As nearly as possible, the general social situation was the same as that which had

obtained in the Introductory Series A and B. In view of the fact that one possible hypothesis to account for the negative deviation depended upon a critical or doubtful attitude on the part of the subject toward the precognition hypothesis, no effort was made to render this hypothesis (precognition) plausible--that is, if the adult subject possessed doubt about the possibility of precognition, that mental state was not disturbed. This preservation of attitudes is difficult, and there is no measure of the success actually attained. It can only be said that whereas with the child subjects considerable thought and attention were given to making the tests as interesting and plausible as possible, little was done in this direction with the adults. On the other hand, nothing was deliberately done to annoy or disturb the subjects nor to increase their scepticism about precognition. The aim was to maintain the *status quo*.

Group effects. No hard and fast rule was laid down regarding the number of subjects that would be allowed to participate at a time. At first it was customary to allow one subject to work at a time. Later they were allowed to work in pairs, and still later the adults were allowed to work in groups. Most subjects enjoyed working in pairs and evidently enjoyed still more working with a greater number participating at a time. As the effect of greater numbers was observed, this way of working became more popular with the experimenters as well. (See below "Group test effects, etc.")

Reward vs. no-reward series. In order to test the impression given in Introductory Series B that offering rewards for high scores might have accentuated the negative deviation during the latter part of that series, it was decided to divide each experimental round on the table into a reward and a non-reward half. That is, four runs on one side of the table were to be subject to the reward schedule and the four runs on the opposite side were to be unrewarded. (In order to remind the subjects of these differences, since some of them tended to forget it, signs were put on the table to indicate which runs were for reward.) This arrangement was continued through five of the seven series.

No-reward controls. But because of the ambiguity of the results in the reward vs. no-reward comparison, it was found necessary also to conduct control series consisting entirely of non-reward tests for each of the two Divisions. In this part of the experiment, it was explained to the subjects at the start that no rewards would be given on either side of the table, no matter what the scores turned out to be.

Nature of rewards offered. In the first few runs of the first series, the rewards were small, consisting of candy and movie tickets. A schedule of money rewards was then introduced according to the following schedule:

Score	Reward
10	$ .50
11-12	1.00
13-14	5.00
15-17	10.00
18-22	25.00
23-25	100.00

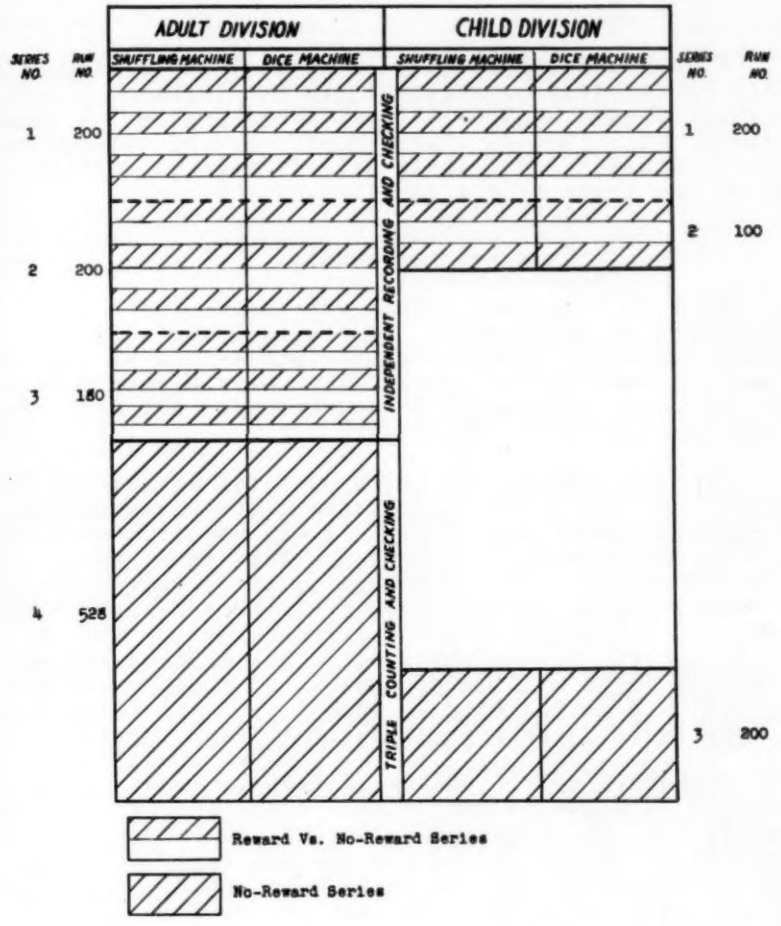
In addition, small rewards in the form of candy bars were given for scores of 8 and 9. Subjects were informed of the reward schedule in advance of the test and were promptly given their winnings as they were earned.

Evaluative Methods

For the computation of the probability of the occurrence of the deviations obtained--that is, deviations from the theoretical mean chance expectation--the critical ratio method based on the standard deviation of the binomial ($\sqrt{npq}$) has been used throughout. In addition, the critical ratio has been obtained from the empirically determined standard deviation. The critical ratio of the difference between Child and Adult Divisions has likewise been obtained, using both the theoretical and the empirical standard deviations. Further, the chi-square values of score frequency distribution have been determined for each of the two large divisions. Also, the combination of critical ratios of sub-series by a chi-square procedure (11: p. 386) is followed in evaluating the results. The contingency method would recognizedly represent a further method of use in the same end, but in the absence of special need and in view of the enormous amount of work involved for so large a mass of data, it has not yet been applied in this instance.

Cross-checks were made on five out of the seven series--that is, on all upon which written records were available. The target series was checked against the corresponding key card of the *next succeeding* round of runs, thus providing an empirical control series, affording a basis for (a) determining whether a chance score distribution obtains: (b) checking on the question of recurrent target card patterns of any origin; (c) discovering effects of possible inadequacy of shuffling; (d) checking on possible sensory cues or peeping on the part of the subjects in

Figure 3
DIAGRAM OF THE EXPERIMENT



For Results See Table IV

the placing of the target cards; and (e) detecting undue occurrence of peculiarities of distribution in the key cards. In dealing with some of the problems just mentioned, the results from the cross-check are evaluated by critical ratio (both theoretical and empirical) and chi-square methods.

Finally, methods are used to deal especially with corrections required or suggested in the light or possible counter-hypotheses such as the matching hypothesis, stacking effect, the optional stopping effect, and various selection hypotheses. All these methods will be further discussed in the light of the specific results.

Experimental Results

In both Divisions, there were conducted a total of 40,200 trials or 1,608 runs. Of these, 1,108 runs made up the Adult Division, distributed in four series of 200, 200, 180, and 528 runs respectively. The Child Division contained 500 runs, made up of three series of 200, 100, and 200 runs respectively. (See Figure 3.)

Adult Division. In terms of total deviations for the divisions, the 1,108 runs of the Adult Division gave a total negative deviation of 239; it was negative as anticipated from the Introductory Series. This is an average deviation per run of -.22, representing an average score of 4.78. The critical ratio based on the theoretical standard deviation is 3.59; and on the empirical, 3.44. The chi-square of the fit of the observed frequency to a binomial distribution is 34 with 12 degrees of freedom, which is equivalent to a probability of .00035. By all three measures, the results are significantly different from chance.

Child Division. The Child Division gave, as was anticipated, a positive deviation; it was slightly larger in average than the negative one of the Adult Division, 500 runs yielding 123 hits above the theoretical mean. The average deviation of .25 (an average score of 5.25) gives a critical ratio of 2.75 by the use of the theoretical standard deviation and 2.63 by the empirical. The chi-square evaluation of the frequency distribution of the scores gives 21 with 10 degrees of freedom, which is equivalent to a probability of .012. These measures, too, are regarded as significant, except the last given; this is .002 above the criterion. (See Table I).

Difference between the Divisions. The critical ratio of the difference between the means, based upon the

observed standard deviations, of the Adult and Child Divisions is 4.10. The critical ratio of the difference based upon the theoretical standard deviation is 4.29.

In view of these various significant criteria, the operation of extra-chance factors is clearly indicated. It is in order, then, to give consideration to other hypotheses--first to those which may be regarded as alternative to the experimental hypothesis under investigation (i.e., precognition) and later to hypotheses of positive bearing that may be indicated by the results.

Table I.*

COMPARISON OF GROSS RESULTS

ADULT AND CHILD DIVISIONS

	Runs	Aver. Dev.	Total Dev.	Theor. C.R.	Emp. C.R.	Prob. Score	χ^2 Freq.
Adult Division	1108	-.216	-239	3.59	3.44	.00035	
Child Division	500	+.246	+123	2.75	2.63	.012	
Diff. between Adult & Child Division		.462		4.29	4.10		
Pooled[3] Adult and Child Division	1608					.0107	

* The original scores of all series reported on in this paper will be furnished interested readers in mimeographed form upon request as will be, also, results of all cross-checks.

3. This pooling ignores the fact that opposite signs of deviation were the experimental objectives in these two divisions. This figure is given, nevertheless, for whatever interest it may have. If the total deviations of the Adult and Child Divisions are pooled regardless of signs, on the ground of there having been opposite sign objectives, the total deviation is 362; the theoretical critical ratio, 4.51; and the empirical, 4.33. Since each Division is independently significant, this pooling adds nothing and is included for those who accept the two Divisions as dealing with the same basic process.

Counter Hypotheses

Counting errors. Of the Adult Division, the first three series totalling 580 runs were conducted with independent recording; that is, without reliance on the triple counting procedure. These 580 runs gave a total negative deviation of 161 (average score, 4.72) which yields a critical ratio of 3.34. (See Table II.) The same critical ratio is given by the empirical method, also. Similarly in the Child Division 300 runs were conducted with independent recording. These gave a positive

Table II.

COMPARISON OF ESP SCORES* AND CROSS-CHECK SCORES

	Runs	Dev.	Theor. C.R. or χ	Emp. C.R.	Prob.# χ (C.R.)	Prob. χ^2 Score Freq.	Prob. χ^2-Comb. C.R.'s
Adult Div.: Series 1-3	580	-161	3.34	3.34	.00042	.02	.0068
Cross Check Series 1-3	580	- 62	1.29	1.30	.0985	.71	
Child Div.: Series 1-2	300	+ 98	2.83	2.75	.00233	.03	.018
Cross Check Series 1-2	300	- 69	1.99	2.05	.0233	.58	—
Diff. Adult and Child	—	(av. dev.) .604	4.25	4.16	.000011	—	—
Diff. Cross Check	—	(av. dev.) .123	0.86	0.88	.195	—	—

* Only those are included on which cross check data are available: Series 1-3 of the Adult Division and Series 1-2 of the Child Division.

\# Multiply these probability values by 2 to make allowance for the possibility that had deviation signs not been as experimentally anticipated, they would have been recognized (i.e., as other than 1).

deviation of 98 with a critical ratio of 2.75. Accordingly, the results may be taken either in their entirety, including the triple counting series (i.e., the last series in each Division), or without, without seriously altering the statistical significance of the report as a whole.

Errors of recording, computation, clerical work, etc. Since recording errors could not under the circumstances of the procedure described above operate to affect

the critical ratios obtained, no further consideration will be given to errors in this general classification. Not only has the clerical work been independently checked, but likewise the arithmetical computation.

Sensory cues. If the subject obtained cues from the cards, or for that matter, looked at them outright, he could have at most grouped the cards of one symbol on one particular key card. This would in the extreme have led to the placing of five circles opposite one key card space, five squares opposite another, etc. While this did not happen in the extreme or it would, of course, have been detected at once, there is the possibility that it happened to some real but unnoticed extent and that there was as a result more grouping of symbols than would have occurred by chance, or than was explainable as due to ESP ability. To differentiate between such an effect due to sensory cues (or other spurious factors) and one due to ESP matching of the key cards, it is necessary only to reflect that if the latter is the determinant, it would presumably lead to more hits (or to fewer-than-chance hits if systematic avoidance occurred)--that is, the groupings would be opposite the right key cards in the check-up; whereas if due to spurious factors, there would be no such grouping except by accident. In that case, the average in the long run would not be affected. Such indeed is the case as revealed by the comparison of the actual deviation of the ESP scores with the deviations obtained from the cross-check scores (in which cards were matched with key cards they were not intended to match). Both in the Adult and Child Divisions and in both methods of evaluation, the critical ratio (both empirical and theoretical) and chi-square, in all a total of six separate evaluations, none of the cross-check results were found to be significant. Nor for that matter were even the individual series themselves, representing, in all, ten independent analyses. This is in contrast to the various criteria of significance rendered by the ESP results themselves. For comparison of ESP scores and cross-check scores, see Table II.

The question, however, of whether an effect from sensory cues leading to a grouping of cards might not affect the results even though the average score itself was not affected,is a legitimate one. Accidental coincidences of grouped symbols would in some instances raise scores and in others lower them. Thus an unusual distribution of scores might be obtained even though it would average approximately five. This might contribute a significant chi-square value. The fact, however, that the cross-check did not yield significant chi-square values indicates that

such grouping, independent of ESP grouping, did not occur, at least appreciably, and accordingly the significant chi-square produced by the ESP scores cannot be

Table III

SUMMARY OF ALL MECHANICAL SELECTION PRECOGNITION
Tests Completed to June, 1939.* (Two Experimenters)

Experimenters	Runs	Dev.	S.D.	Theor. C.R.	χ^2	Prob. χ^2 Comb. C.R.'s
Stuart & Pratt	1368	+98	74.0	1.32	1.74	
Crampton & Smith	813	-69	57.0	1.21	1.46	
Rhine & Gibson (Adult)	1108	-239	66.6	3.59	12.89	
Rhine & Gibson (Child)	500	+123	44.7	2.75	7.56	
Total	3789				23.65	.00009

* For series incomplete at that time, see Table VI.

attributed to the hypothesis of grouping due to sensory cues.

It might further be questioned whether the usual standard deviation could be properly applied to the results if there is a possibility that through sensory cues, or other non-experimental factors, the spurious groupings in question might have been produced, even on the supposition of a purely chance series. But the same evidence which met the objection considered in the paragraph above applies here likewise--that is, if there is no evidence of the supposed grouping attributable to sensory cues (and like factors) as is shown by the fact that the empirical C.R. and the chi-square of the cross-check score-distribution are insignificant (See Table II), it indicates that the departure from the binomial produced by the ESP scores was not attributable to unusual groupings by such improper factors as sensory cues.

The stacking effect, card-warping effect, inadequate shuffling hypotheses. The stacking effect is simply a name for the general possibility that because of one circumstance or another some non-random grouping of the target cards might occur, producing a stacking of cards of one symbol opposite a given key card space without the arrangement having had anything to do with precognitive ESP ability. Peculiar warping effects of the

Table IV

RESULTS OF THE EXPERIMENTS REPORTED, AS DISTRIBUTED OVER VARIOUS SUBDIVISIONS

Subdivision	Method of Selection	Reward or No-Reward	Runs	Dev.	Aver. Dev.	S.D.	C.R.	χ^2 or C.R.^{2}
Adult Div.	Shuff.	Rew.	50*	- 17				
Series 1	Mach.	No-R.	50	- 7				
Rew. vs. No-R.	Dice	Rew	50	- 10				
	Mach.	No-R.	50	- 3				
Series Total			200	- 37	.185	28.3	1.31	1.72
	Shuff.	Rew.	50	- 9				
Series 2	Mach.	No-R.	50	- 22				
Rew. vs. No-R.	Dice	Rew.	50	- 23				
	Mach.	No-R.	50	- 23				
Series Total			200	- 77	.385	28.3	2.72	7.40
	Shuff.	Rew.	45	- 19				
Series 3	Mach.	No-R.	45	- 25				
Rew. vs. No-R.	Dice	Rew.	45	+ 12				
	Mach.	No-R.	45	- 15				
Series Total			180	- 47	.261	26.8	1.75	3.06
	Shuff.	No-R.	132	- 11				
Series 4	Mach.	No-R.	132	+ 17				
No-Reward	Dice	No-R.	132	- 49				
Control	Mach.	No-R.	132	- 35				
Series Total			528	- 78	.148	46.0	1.69	2.86
Adult Division Total			1108	-239	.216	66.6	3.59	12.88
			$P(\chi^2 = 15.04) = .0046$					
Child Div.	Shuff.	Rew.	50	+ 40				
Series 1	Mach.	No-R.	50	+ 11				
Rew. vs. No-R.	Dice	Rew.	50	+ 1				
	Mach.	No-R.	50	+ 15				
Series Total			200	+ 67	.335	28.3	2.37	5.62
	Shuff.	Rew.	25	+ 18				
Series 2	Mach.	No-R.	25	+ 7				
Rew. vs. No-R.	Dice	Rew.	25	+ 6				
	Mach.	No-R.	25	0				
Series Total			100	+ 31	.310	20.0	1.55	2.40
	Shuff.	No-R.	50	- 8				
Series 3	Mach.	No-R.	50	+ 6				
No-Reward	Dice	No-R.	50	+ 29				
Control	Mach.	No-R.	50	- 2				
Series Total			200	25	.125	28.3	0.88	0.77
Child Division Total			500	123	.246	44.7	2.75	7.56
			$P(\chi^2 = 8.79) = .032$					
GRAND TOTAL			1608					

(Combined by Series) $P(\chi^2 = 23.83) = .0012$
(Combined by Div.) $P(\chi^2 = 20.44) = .000037$

* Groups of Runs in this column also represent columns in the record book and different positions of runs on one side of the experimental table.

cards leading to tendencies for certain cards to cling together, others to separate more easily, are sometimes thought of as possibly contributing to such an effect. Inadequacy of shuffling, combined with habits on the part of the subject of placing cards in a certain pattern might likewise be suspected on superficial consideration to contribute to some such result. But from whatever hypothetical source the case of such grouping may be supposed to come, including the more plausible source discussed above under sensory cues, the methods of checking up on the hypothesis would be essentially the same; namely, the analysis of the cross-check data. As already shown, the evidence from the analyses in the present instance (and in all earlier work similarly examined) (11: p. 43) gives no grounds for this hypothesis.

Selection of data. Certain aspects of the general question of selection of data may be answered easily

Table V

EFFECTS OF REWARD UPON PERFORMANCE

Subdivision	Condition	Runs	Dev.	Av. Dev.
	Adult Division			
Series 1	Reward	100	-27	
	No-Reward	100	-10	
Series 2	Reward	100	-32	
	No-Reward	100	-45	
Series 3	Reward	90	-7	
	No-Reward	90	-40	
Series 1-3	Reward	290	-66	-.23
	No-Reward	290	-95	-.33
Totals Reward vs. No-Reward		580	-161	-.28
Series 4	Control No-Reward	528	-78	-.15
	Child Division			
Series 1	Reward	100	+41	
	No-Reward	100	+26	
Series 2	Reward	50	+24	
	No-Reward	50	+ 7	
Series 1-2	Reward	150	+65	+.43
	No-Reward	150	+33	+.22
Totals Reward vs. No-Reward		300	+98	+.33
Series 4	Control No-Reward	200	+25	+.13

and objectively; others may be answered equally objectively but require elaboration. So far as the selection of data through loss, discarding, or deliberate manipulation of grouping is concerned, the answer on the records (and in duplicate) is that there was none. The series reported herein were announced in advance and official record sheets provided for the purpose. There are at least two independent and complete score records on file for all the work described. In resume: from the beginning of the investigation herein reported, on December 8, 1938, to the end of the series on March 17, 1939, no single score has been omitted from experiments of the character described, conducted in this Laboratory.

The more difficult aspect of the selection hypothesis concerns the propriety of the presentation of the particular group of data which is reported in this paper. The principles followed in this instance are those believed to be commonly in practice among experimental psychologists and experimenters in general. The radical implications of the precognition hypothesis, however, make it imperative to consider afresh all such principles and to take note of any possibly unsafe assumptions that may be involved.

First, it is general practice for an investigator in any field to regard his own work as an adequate sample for separate statistical treatment and to evaluate it independently of any similar work produced by others. Second, it is likewise common procedure that a given set of conditions be sufficient criteria for the delimitation of a sample of statistical data. If the sample treated is representative of the work done under the specified conditions, work done under other conditions, however small the difference, may be excluded from the evaluation so long as there is any possibility that the difference in condition might conceivably contribute to determination of the results. Third, there is commonly no question raised over the policy of the experimenter in deciding upon the point at which he issues a report of results--that is, the stopping point, temporary or otherwise, in the general exploration which he is undertaking. Aside from certain niceties expected in the way of rounding out a series of "sufficient length" there are no fixed standards of procedure governing the termination of a research program. The research may be mapped off in blocks or units of specified length, but whether or not the units adopted will be repeated at a given point in the series is generally left without question to the judgment of the experimenter.

But in view of the special concern over methodology in this research, special consideration will be given to the justification of the selection made in this case. The first question might be put as follows: On what grounds do I restrict the report to cover my own work and exclude from present consideration work of comparable character by other investigators? The answer is that in this instance, as with most psychological experiments in which the experimenter is part of the environment, the work by a given experimenter may very well involve a difference in environmental determinants of the results. It is important in the very interest of comparison of experimenters to have the work of a given experimenter separately treated and presented. There is, however, a secondary consideration which should be kept in view at the same time; namely that if there are a number of experimenters using comparable conditions in the investigation of a given hypothesis, the general results have a bearing--that is, if a probability is found for the results obtained by one of the several experimenters, it is relevant to take note of the number who have been making the attempt. A corrective factor for such a situation is well known [i.e., $p' = 1-(1-p)^n$].

It would appear simple enough to apply the corrective factor just mentioned to the case in hand if everyone were to agree as to just what would constitute effectively similar procedures. Making liberal allowance and overlooking differences that might well justify discrimination (e.g., very different types of subjects were used), the work of two other experimental series completed but as yet unpublished (one undergoing analysis for early publication) might be regarded as comparable and included in the evaluation. I refer to research on precognitive matching conducted during the same year by Mr. C.E. Stuart and Dr. J.G. Pratt of this Laboratory, using the mechanical shuffler, and to that of Mr. Burke M. Smith and Miss Barbara Crampton, using a similar procedure. The results of these two series are given in summary form in Table III, along with the results of the two divisions of the present report and a critical ratio is presented for each, with a combination of these critical ratios by a chi-square method suggested for this purpose by Greenwood. It will be seen that while neither of the other two series produces a significant critical ratio[4], they

4. It should be mentioned that the forthcoming report of the research by Stuart and Pratt will show that in spite of the insignificant critical ratio important relations have been found upon closer analysis.

do not affect the significance of the results of the three groups of research taken as a whole. ($p = .00009$)

The next question that arises is: Are the grounds adequate for restricting this report to work with mechanical selection of cards? In my previous report, I gave a complete account of the research on pre-shuffle card calling in which the shuffling was manual. It was later pointed out that another method of shuffling was called for to eliminate the possibility of the ESP shuffle (10). It is *necessary*, then, to separate the mechanical selection work from the earlier report. However, other methods of selection have been used; for example, the tests with the use of hand thrown dice mentioned in the Introductory Series A and B.[5] Later work has been done with the method of rearranging the cards involving specified sets of numbers to be obtained from a newspaper of specified future date. These differences in determining the card order are sufficiently great to warrant separate treatment of the results. However, if there be any disposition to question the propriety of this separate treatment, chi-square combination of the critical ratios is again possible. Table VI provides the chi-square in question. It will be seen that the data are there assembled from a considerable variety of conditions and the critical ratios of the different groupings are combined so as to show the effect of grouping upon the estimate of the probability of the occurrence of the C.R.'s. The probability is $10^{-\frac{1}{2}}$.

A similar question might be raised with regard to the restriction of the research herein reported to that conducted with two or more experimenters present, each furnishing a check upon the other. In answer to this, it would have to be said that such a separate treatment as the present report affords would, as a matter of fact, be demanded by most investigators in the field at the present time.

The chi-square method of combining critical ratios is advantageous in the evaluation of this work as a whole in view of the varied conditions; e.g., some researches

5. The special question might occur as to the ground for excluding the Introductory Series A and B. The reasons are, in brief, that the procedures and conditions were quite different from the Child and Adult Divisions. To cite only one difference: the selection of cards was all done with dice, and these were all hand-thrown. Mechanical selection did not occur in either series. But if these series (A and B) were pooled with the Adult Division, the C.R. would still be of almost equal significance: 3.43 as against 3.59.

were conducted with the express goal of obtaining negative deviations, as in the case of the Adult Division of the work reported above. The work done in the correspondence tests for precognition is likewise negative. However, the use of the chi-square combination method referred to above makes allowance for the possibility of deviation in either direction in every series.

The principal question remaining for discussion under the selection hypothesis is that dealing with the legitimate determination of the stopping point. As Greenwood has pointed out (4), practically every statistical report potentially involves optional stopping. There is no way known as yet to insure with finality against some option as to whether the experimenter will continue or will stop (e.g., to make a report). He may, as was done in the work herein reported, state at the outset that he will do a stated number of runs. When that is through he may repeat and again the same possibility is left. Stating in advance that he will not repeat nor continue under any circumstances or with any variation would, of course, remain a possible solution and would reduce the question to one only of the ethics of the experimenter in adhering to his own self-limitation. Since, however, no experimenter probably wishes to commit himself against further investigation, the corrections arrived at by Greenwood offer the best measure at present available of security against possibility of error on that score in the interpretation of results. This correction represents an upper limit. When the Greenwood correction is applied to the probability obtained from the maximum critical ratio produced by the work reported in this paper, it is increased from .000,058 to .000,464 and still remains significant.[6]

From the point of view of practical empirical observation, the optional stopping hypothesis stands upon a very weak case in any event. In Greenwood's entire series of 500,000 matchings of cards against calls he examined the possibilities of finding a favorable point of stopping and found in the entire series no point with a critical ratio of the cumulative deviation as great as that obtained in the present series. Optional stopping, then, while theoretically important, is here practically of minimal concern.

---

6. I am indebted to Dr. Greenwood for this computation. His method of determining this upper bound is in the course of preparation for publication.

There is a consistency of trend to the deviations in both the Child and Adult series that is of considerable reassurance also against the hypothesis that the favorable deviation critical ratios might be due simply to finding favorable stopping points. As may be seen from inspection of Table IV, the natural grouping of the results gives predominantly negative deviations for the Adult Division and predominantly positive deviations for the Child. In 23 of the 28 logical sub-divisions, the deviation is that which would have been experimentally anticipated. The probability of this occurring in a chance series is .000,46[7] [obtained from the expansion of $(p + q)^n$]. Neither this significant distribution of signs of deviation nor the chi-square of the frequency distribution of scores is wholly independent of the deviation critical ratio evaluations, yet since it is possible that they could have been insignificant while a critical ratio of the total scores was significant, they do offer further support against the applicability of the optional stopping hypothesis.

This lengthy discussion of the optional stopping hypothesis is regarded as fully warranted in view of the possible applicability of this hypothesis to all analyses of ESP results (and, of course, to all other statistical findings). Until the fullest implications of Greenwood's analysis of the optional stopping hypothesis (4) are worked out, it will be necessary to be guarded in drawing conclusions from marginal critical ratios (e.g., in the region of 2.5 to 3.0). It is safe to say the proper pooling of results eliminates the possibility of an optional stopping effect up to the end of the final series included in the pool; that is, there is only the option as to the pooling point. The same cannot be said of the combination of critical ratios by the chi square method, in that each series treated is subject to the optional stopping hypothesis. While, therefore, the combination chi squares of Tables III and VI are subject to this counter-hypothesis, any serious concern may be removed by calling attention to the fact (1) that this method of evaluation is not the one primarily used and hence could not have been the basis of determining the pooling-point. (2) The method and its results could be ignored without affecting the conclusions. It is secondary and supplementary; though it is not a wholly independent measure.

---

7. To allow for the possibility that, had the signs gone in the direction opposite to that experimentally expected, the results would still have been considered statistically, this value is multiplied by 2, giving .000,92.

This is as far as it seems warranted to go for the present into a consideration of the optional stopping hypothesis of selection of data. *Perhaps the best assurance available on this question is the fact that pooling at any point since the beginning series in precognition tests would have given a significant C.R.*

It has already been stated that the regular procedure followed required the statement, in advance of an experimental series, of the number of runs to be conducted; upon which the Record Librarian issued the required number of serially numbered officially stamped sheets. These sheets were all to be returned to him and thus accounted for, whether filled out or not. In 5 out of the 7 series, this was done exactly as was intended, with all the record sheets filled out. In the third of the five series of the Adult Division, however, the quota was completed for all but 20 runs out of the 200 prescribed when the Christmas holiday season arrived. Rather than hold the series incomplete for this small number of runs and to resume with the subjects in a possibly very much altered attitude, the series was declared closed and the record sheets returned blank. In the fourth series of the Adult Division, there had been scheduled in advance a total of 400 runs. At the close of this series, which was to be the last, an opportunity provided itself to introduce a group of adult visitors to the Laboratory and because of growing interest in apparent group effects on scoring an emergency extension of the series was requested and record sheets provided to extend the series from 400 up to the total of 580 which had been the total of the three preceding series on reward vs. no reward. As it happened, with the end of the session of group tests 128 runs had been added, making a total of 528. With no comparable subject situation in prospect for the completion of the series and with other experiments occupying the time and interest of both subjects and experimenters, the remaining unused record blanks were at length turned in to the Librarian. The slight irregularity or departure from the prescribed schedule in this instance may be dealt with in either or two ways. First, the extension series of 128 might be set aside, if there were any warrant to do so, without affecting the general significance of the report, or it may be endorsed as a legitimate part of series 4. The latter treatment has been followed as the simpler procedure, since the critical ratio of the Adult Division is altered only from 3.40 to 3.59 by the inclusion.

Hypotheses of competence and trustworthiness. The experimental procedures for all of the work herein

reported was so planned as to make two experimenters mutually responsible for the safety of the results. In order to avoid clerical errors of any kind, each experimenter kept his own records and kept these in his possession until totals were computed and checked. Even then a record of the totals was left in the possession of each experimenter. Consequently, there was no moment at any time when one experimenter could make an error without a discrepancy occurring between the two records and requiring checking up and eventual discovery. The rule requiring the acceptance of the score nearer the chance mean in the event of discrepancies safeguarded the experiment fully from errors of this nature.

Improper arrangement of data. Under this heading might be proposed the question whether the Child and Adult Divisions scores should not have been pooled instead of treated separately. Thus the deviations would tend to cancel and the results would give an insignificant critical ratio. Even so, the score frequency distribution would still give a significant chi square; P = .0107. [See Table I]. The grounds for keeping the divisions separate in the evaluation of the results may be reviewed briefly. First, the experimental divisions were planned separately and indicated as separate experiments in the advance descriptions filed at the start.[8] Separate batches of record sheets with different serial numbering were issued for the two divisions. Moreover, all the work of a comparable character done previously with children under similar conditions had led to the expectation of positive deviations. On the other hand, the previous work of comparable character conducted with the adult subjects used in the series reported here had given negative deviations, as the Introductory Series A and B, and a continuation under the circumstances would hardly be construed as having the same expectation as that of the Child experiment. Furthermore, the repetition of series after series, each Child series yielding positive results and each Adult series yielding negative deviations, could hardly make more plain and irrefutable the direction of expectation and the purpose of the comparison of the Child and Adult performances. Finally, it is possible to apply a method of evaluation, a chi square combination of critical ratios, which can be used with

8. The *direction* of deviation expected was not given in the advance statement of experimental plan in any of the seven series; these statements were open to inspection by the Laboratory staff-members, most of whom were serving as subjects. For them to know that a negative trend was expected of them might have changed the actual trend.

disregard of the direction of deviation. Thus allowing that each series might equally well have been expected to be positive or negative, the critical ratios are combined and the combination chi square obtained. As shown at the bottom of Table IV, the result is still significant; (P = .0012). Similarly the critical ratios of the two Divisions may be combined, as is done in the same table (P = .000037).

Alternatives to the Precognition Hypothesis that Involve ESP

If, as would appear from the evidence available, the results cannot well be attributed to chance and since the experimental conditions could hardly have permitted sensory cues, card defects, shuffling inadequacies, errors in recording or other clerical work, or the selection of data to have influenced the results in favor of high scores, it is necessary to consider what other hypotheses are alternative to precognition. It appears necessary to assume that some causal relation existed between the performance of the subject and the extra-chance results, and that in some way the subject was effectively oriented in his responses. Thus he was behaving as if he were able to discriminate between situations, one of which (the order of the key cards) did not at the time (in our accepted usage of the term) exist. In laying his cards opposite five blank key spaces to be occupied later by the key cards when they were selected, he was, it would seem, responding to a purely future situation--reacting reliably to a seeming unreality. Superficially, at least, this would seem to be impossible, perhaps as completely absurd to present conceptions as anything which could well be fancied.

But the alternatives of which I am aware are hardly less so, unless one rebels against the inference of causality from the statistical results of an experiment (and even here common sense is allied with statistics, as for example, to cite only one instance, in all seven series and in all 14 subseries of Table V, the sign of the deviation is as expected. In other words, the student in his discomfort over the alternatives may choose to await more statistical evidence of the same kind, but hardly to find fault with the quality of that presented in this report.

The two most natural alternatives to the inference of the hypothesis of precognition are rational inference and exercise of control over the precognized series of

events, making it occur according to prediction. To follow the former alternative in the present instance would mean to develop the hypothesis that, as the subject matches the cards opposite the blank spaces, he is combining his ESP impressions of the cards with a (subconscious) rational set of inferences as to the probable occurrence of the pattern of key cards to follow. With our present knowledge, however, it would seem entirely gratuitous to assume any basis for rational inference as to the probable appearance of the sequences of key cards either by attempting to judge in advance the probable rearrangement of the cards in the shuffling machine, or to attempt reliable inference as to the most likely total number of spots on the upturned faces of the dice. No hope for escape from the precognition hypothesis over this route seems, to the present writer at least, to be offered.

The other alternative requires that the subject, or some other interested party, to some extent determine the supposedly random events--in this case the selection of a given pattern of key cards. The person most likely to play such a role would be the experimenter whose task is to stop the shuffling machine and dice machine. He could by ESP have some (of course unconscious) impression of the order of the cards laid down by the subject, even though he is located in a different room, and he might also, it is supposed, use this ability to good effect in apprehending the contents of the shuffling machine and could decide upon a favorable moment to stop it. He was, however, required by the instructions to stop it immediately on hearing the buzzer (each machine had to be stopped in a given position); i.e., at the next suitable stopping point. This meant that no further revolution was permitted and that accordingly the precise stopping point was determined with a minimum of latitude left to the experimenter. He was barred from exercising any control over the machines that could determine the order of the cards. If the subject is supposed to have some knowledge through ESP of the behavior of the cards in the shuffling machine, the constant motion of the machine right up to the point where he lays down his last card must prevent any use of this supposed knowledge, unless it be of a precognitive character, from being utilized in the distribution of the deck.

What is left, then, as the only possibility at present conceivable is that there is some way by which the experimenter or subject can influence the behavior of the cards in the shuffling machine and the dice in the dice machine so as to make them favor a pattern which

will give a better-than-chance score when the cards are checked against them. It will perhaps seem a no more preposterous supposition to imagine some extension beyond the musculature of the influence of mental processes, something corresponding to motor activity (extra-motor behavior), as extra-sensory perception does to sensory. As a matter of fact, neither this hypothetical kinetic effect of mental process ("telekinesis" in the terminology of psychic research) nor the supposition of extra-sensory perception of card orders that do not yet exist are easy to entertain in relation to present day conceptions of the nature of mind.

If one is to make his choice purely on the basis of such logical grounds as there are, he will doubtless find it easier to prefer the precognition hypothesis to the kinetic alternative described. The increasingly extensive background of the ESP research itself favors this choice. It has also been pointed out above that since the evidence against the spatial limitation of ESP is considerable (11: ch. XIII) and since spatial and temporal criteria are apparently interdependent, it is logical to expect it to be free from time restriction. In the light of this and other considerations, the critical reader, faced with this dilemma and lacking experimental evidence for the "telekinesis" alternative, will presumably favor precognition. There is, however, no need to reach a final conclusion upon the evidence reported in this article. Others will follow as a matter of course, from other experimenters.

Some Incidental Findings

The effect of rewards. In the five series, three of the Adult and two of the Child Division, in which each experimental round of eight runs was divided equally between reward and no-reward conditions, the results do not present a reliable difference. With the children, the deviation for the reward runs was approximately twice that of the non-reward runs; +65 for 150 runs for reward, and +33 for the same number of non-reward runs. (See Table V) It was obvious to the experimenters that the children were much more interested in the reward runs and were frank in their expressions of regret if they obtained a high score on the no-reward runs. They did not, however, express any desire to score below chance. They understood that the experimenter always wanted high scores and they were willing to try for them. They simply wished that any high scores they might obtain would come in the reward series.

With the adult series the case was somewhat reversed both in scores and in attitude. There was no doubt of the desire of most of the adult subjects to obtain the reward, but since most of the subjects were not only personal friends, but were in some way connected with the research of the Laboratory, there was genuine appreciation for the importance of high scores on the no-reward runs. There was accordingly much more appearance of conscientious effort on the part of the adult subjects to do good work on the no-reward runs as well. There is, however, no criterion of the impressions just mentioned other than the observer's personal judgment. The actual deviations obtained in the three series for adults were for 290 reward runs -66 and for the same number of no-reward runs -95.

It appears that an uncomplicated no-reward control was desirable with a view to providing a basis for comparison for the mixed (i.e., reward vs. no-reward) series. As nearly as possible the same subjects were used in approximately the same number of runs in the control series. This could not, however, be done exactly. Two hundred runs were required of the child subjects and 528 of the adults. As would be expected, the motivation--particularly with the child subjects--was noticeably different so far as can be judged by expressions of interest. There were not, however, any expressions of negativism from either group of subjects. There were in both Divisions, but particularly with the children, definite expressions of preference for the reward vs. no reward series as distinguished from the no-reward series.

The result for the children was a positive deviation of 25 in the 200 runs. Compared to the reward vs. no-reward series which altogether averaged .32 per run above the mean for 300 runs, this control series averaged only .13 above. Taking in order first the reward runs and second the non-reward runs appearing in the reward series and third the purely non-reward series, we have the following average deviation per run: .43, .22, .13. While these differences are not significant, they are sufficiently in line with general expectation to be at least suggestive. With the children at least, the rewards would appear to be generally associated with higher scoring. No-reward runs occurring in the same experimental round with reward runs might be said to catch something of the advantage attaching to the rewards, while the purely non-reward series though still positive in trend noticeably suffered from reduced motivation. It is repeated that these are regarded as suggested but not proved by the results.

In the adult no-reward series, likewise, the deviation was considerably less; whereas for the 580 of the mixed reward and no-reward series the average deviation had been .28, it dropped to .15 for the 528 runs of the control series. The sequence for reward runs, no-reward runs associated with reward runs, and the control no-reward series is as follows (giving average deviation per run): .23, .33, .15. Table V presents the results on the comparison just reviewed. (See also Table IV for the subdivisions of the series from which the comparisons are made.)

The most general and consistent effect suggested by the reward study is that merely having rewards associated with the experiment (as in the reward vs. no-reward series) makes the tests more interesting, the average deviation greater. Not only was this true of the reward runs but to a greater or less degree of the non-reward runs made during the experimental sessions when rewards were given. This spread of effect is somewhat similar to the findings of Woodruff and George (15). They obtained a significantly higher average score for the entire month during which small prizes were offered in alternate weeks to the highest scoring subject of the week; but they did not find that there was any difference between the no-prize weeks and the prize weeks for the month in which the prizes were given. It appears that a very diffused improvement of interest attached to the experiment through the month in which rewards were offered. Though significant differences are lacking in the present work, it is not difficult to concede some enhancement of deviation as due to the reward, but it appears that the reward effect is somewhat generally diffused to cover the entire experimental session, especially with the adult subjects; much the greater differentiation between reward and no-reward runs occurred with child subjects.

It is relevant to inquire whether the lower averages of the control (no-reward) series were not incidentally due to decline trend occurring with the progress of the experiment over a long period, since such effects have been earlier reported (8). Table IV shows, however, that (a) in the Child Division there was little decline from series 1 to series 2; and that the drop in series 3 was accordingly abrupt; and (b) that while there was a higher average at series 2 of the Adult Division than in series 3, series 1 was still lower than series 3; also, the division of series 4 into the originally scheduled 400 runs and the extension of 128 runs, shows that the latter gave a higher average than the former. The decline hypothesis, therefore, does not fit the results.

Shuffling machine vs. dice machine in the selection of key cards. At the beginning of the work reported in this article, provision was laid down in the instructions that the first two sets of keys at the west end of the laboratory table should be selected by the shuffling machine while the two at the east end should be selected by the dice throwing machine. No check-up was made as to whether a difference was occurring between the results of the two procedures until the final analysis of the work here reported. It was then observed that a considerable difference obtained in the deviations from the two methods of selecting the key cards, a difference which like the effect of rewards is insignificant in itself, but which is consistent enough for the two divisions to occasion interest. Pooling all adult series together, there are 554 runs using the shuffling machine and an equal number for the dice throwing machine. The first gave a total negative deviation of 159, and the second (dice) a total of 80. That is, the deviation with the shuffling machine is practically twice that of the dice machine. (Details from which these totals are taken are given in Table IV)

The difference was even more marked in the Child Division. There 250 runs with the shuffling machine give a total positive deviation of 97 while that with the dice machine totals 26. The fact that the shuffling machine is associated with a larger negative deviation in the Adult Division and a larger positive deviation in the Child Division suggests that it is associated with the deviation-producing factors in the experiment--that is, factors of ESP character. But the force of the suggestion is considerably diminished when it is discovered that there is no consistency of trend from series to series--that is, the shuffling machine is associated with higher scoring in one and lower in another, with about equal frequency. To the experimenters it appears plausible to suggest that the differences found in the shuffling and dice selected work may be, if not accidental, more a matter of position and sequence than anything else. The subject usually began at the shuffling end of the table, that is, where the shuffle selected keys were to be placed, followed around the table with the two more dice-selected runs and ended with the last two runs opposite the shuffle selected keys. The "end-effect" hypothesis might well be applicable here, the shuffling machine tending to get whatever advantage attached to the beginning and end runs. It will remain, then, for further investigation to distinguish between this hypothesis and that which supposes that key cards selected by the shuffling machine have a better chance at being associated with high scores than if they are selected by the dice throwing

machine. Obviously if the shuffle and dice selected keys had been reversed at proper intervals in the progress of the investigation, more light would have been thrown upon this point. The question, however, was an incidental one at the time and the need for variation was overlooked.

Group test effects in the work of the Adult Division. At the beginning of the experimental work covered in this report, it was the common practice to have present in the Laboratory only one subject at a time. Later two were allowed to work at the same table and from time to time an occasional trial would be made with a larger number. The maximum that could be allowed to work at a time was eight, since there were eight runs possible to a round. The impression was given that when a number of adults were working together an enhancement of interest occurred and that there seemed to be reflected an increased score effect--that is, larger deviations in the direction experimentally desirable. Accordingly as the program advanced larger portions of the series were given over to the group experiments, except that in the no-reward control series an effort was made to balance the proportion of group tests, making it roughly comparable to that of the other series. A check-up on the comparison of deviations produced with different numbers of subjects present at one time reveals definitely that the impression was a correct one; in fact, the largest deviation occurred with the largest number of subjects. (This effect is confined to the Adult Division since there were only five subjects used in the Child Division all together--not enough to try a round of eight.) From the total of 1108 runs, averaging .216 per run below expectation, the 656 runs done with one to four subjects present averaged .155 below the mean (C.R. = 2.03) while the 452 runs done with eight subjects working at a time or eight or more subjects present in the room averaged .303 per run below (C.R. = 3.22).

Certainly, then, it may be said that under these circumstances *and with the objective of obtaining a negative deviation*, the effect of group testing appears to be a favorable one. Further evidence is required, however, to determine whether this is a general relationship, as the present results do not show a statistically significant difference and the results represent, in part, a *post hoc* selection of data. The principal effect producing the statistical significance of the Adult Division results as a whole comes from work done under conditions of group testing. It must be kept in mind, of course, that in all these instances the subjects were well acquainted with each other, were together for a common purpose, and

all entered into the test procedure with a social spirit resembling that of a game at a party. In comparison to this, the condition of working alone, especially since it involves laying out eight packs of cards at a round, can readily be seen to offer a relatively less attractive situation.

The question of the causation of the different trends of deviation in the Adult and Child Divisions. It may be said at the outset that negative deviations which appear to be of extra-chance character afford considerable difficulty of explanation. Hypotheses have been offered to explain unintended or ostensibly unintended negative deviations, accounting for the effect first as ESP ability plus unconscious negativism, and second as a kind of "bad aim" or ESP illusion, implying no change of motivation but the adoption of a consistently wrong method of exercising ESP so that its function is systematically diverted.[9] In view of the wholly unconscious character of the essential activity involved in extra-sensory perception, it is not difficult to conceive of either hypothesis being applicable. However, the former seems more elaborate since it requires the additional supposition of unconscious and conflicting motivation. The latter, the blind but consistent following of a misleading mental procedure, would at all events seem more applicable to the present situation. Certainly the subjects who participated, adults as much as children, had no reason to be negatively motivated. They were all interested in the subject and such conditions as the introduction of rewards and the group test--conditions that obviously made the test program a more pleasant one for the subjects--accentuated the deviation trend already established. It would seem accordingly that if by strengthening the motivation of the subjects a more negative deviation was obtained, it is likely that it was "bad aim" that was concerned.

I had attempted earlier to account for significant negative deviations obtained with tests involving enclosed cards (sealed opaque envelopes) in the course of which there was considerable delay in informing the subject of the scores, tests mainly conducted by correspondence. I was inclined at the time to look upon this negative deviation as due to the annoyance and frustration experienced by the subject over the prospect of not being able to check up on scores, not to find out what was in the envelope until several days later. This was the negativism

9. These two hypotheses have been facetiously but more graphically called the "double-crossing" and the "crooked barrel" hypotheses.

hypothesis. But it is equally plausible, perhaps more so, to suggest that the subject working in isolation, concentrating intensively on the enclosed cards with a view to sending in his records to the Laboratory, might have been giving enough additional effort that in his unconsciousness of the actual procedure to follow mentally, he--to use a figure--overshot the mark; that is, he seized upon a way of doing a test that systematically reversed or digressed from the casual, normal procedure which ESP would involve. In the present experiment, likewise, the subject being especially cooperative and appreciating perhaps to an excessive degree the significance of the precognition hypothesis, may have kept doing his best introspectively to seize upon a really superior way of exercising his latent ESP ability; he may reasonably enough have laid hold of what was for him a perverse technique and accordingly the more he held to it, the stronger his motivation (e.g., enhanced by rewards), the more he matched the cards in a way that they should *not* have been matched.

Owing to the fact that the tests of the present series were tests of precognition, it is conceivable that the average subject's natural scepticism of precognition (producing feeling of frustration) might account for negative motivation. This was at least faintly suggested by the Introductory Series A in which the clairvoyance half of the series yielded a positive average, and the precognition half gave a negative deviation.

As stated at the outset, there is no easy explanation for the consistent trend of negative deviation by the group of subjects responsible for the Adult Division tests. Fortunately for the bearing it may have upon the precognition hypothesis, this explanation need not for the present be determined. A reliable relationship, even of negative character, would be good enough as a starting point. Nevertheless it remains a puzzling question; and accordingly the remarks that have been made on the subject are not submitted with strong conviction.

Similarity of negative and positive deviations. The average deviation per run of the Adult Division was .22 and that of the Child Division, .25. Since these deviations are of opposite sign, the similarity in average is of interest. The comparison goes further: The average deviation per run for the no-reward series of the Adult Division was .15, and for the Child Division .13; while deviations per run of the reward vs. no-reward series of the Adult Division averaged .28 and of the Child, .33.

Again, the shuffling machine selection gave for the Adult and Child series respectively .29 and .39 average deviation per run, and for the dice machine selections .14 and .10. These approximate similarities are interesting and warrant some consideration as to the factors that may have contributed to them; especially since the similarity would theoretically not be expected.

As was pointed out by Thouless (13) and as is indicated by Foster's ESP quotient formulae (see his article in the preceding issue), if equal positive and negative deviations occur, there is inferable ESP capacity in the two cases in approximately a 4 to 1 ratio. The ESP quotients of the Adult and Child Divisions are respectively 4.40 and 1.25 (computed from the deviations -.22 and +.25, and assuming that the fostered negative deviation of the Adult Division work was equivalent to that of a deliberately low-aim series). This would seem to indicate that whatever is measured by the ESP tests existed in approximately a 4.40-to-1.25 ratio in the two groups of subjects tested under similar conditions.

It must be borne in mind that this similarity of deviations may constitute a random event; but there have been other instances that have some supporting influence upon the case--particularly those in which subjects deliberately tried to obtain negative deviations for comparison with the positive deviations they had obtained in other series. In 1934 I reported the results of high-aim vs. low-aim experiments on two subjects. The first subject, H.P., (8: p. 30) obtained in 11 runs of low-aim, a negative deviation of 3.2 per run. His average for the total of 690 high-aim runs produced by him at the time of the report gave an average positive deviation per run of 3 (although at the time of his low-aim tests the more comparable average positive deviation of 4.2 on 200 high-aim runs represents a closer approximation to the conditions of the low-aim tests). The other subject was Mr. C. E. Stuart, of this Laboratory, (8: p. 70) who conducted his own experiments, serving as subject himself. In 52 runs on which he made both high-aim and low-aim calls on the same card, he obtained an average negative deviation of 1.5 per run and an average positive deviation of 1.6. In the report of Pegram (5) on comparison of high-aim and low-aim ESP tests, it is noted that the low-aim produced a deviation per run of -.65 while the total high-aim work (which was more than five times as great as the low-aim) gave +.30. But when the high-aim average deviation that was obtained during the period in which the low-aim experiments were conducted is considered alone,

the results are fairly comparable; namely, for 665 runs of high aim, an average positive deviation of .56 per run was obtained, and for 1468 runs of low aim, an average negative deviation of .65 per run. Other comparisons are offered by Pegram but those cited would appear to be the more logically comparable. Altogether the data cited raise the question whether the ratio of 4 to 1 referred to above fully covers all the factors inherent in the ESP situation where low-aim and high-aim tests are made (and in the equivalent case represented by the Adult Division in this report in which for some reason a perversion or consistently bad aim of the ESP process operates to produce significant negative deviations when positive scores are sought by the subject).

There are on the other hand at least three instances on record in which negative deviations have been produced which had been relatively small in comparison to the positive deviations with which they might with some reason be compared. First in the case of the subject A.J.L. (8: p. 63), whose work I reported in 1934, the negative deviation obtained (by pressing continuation of the tests against the subject's will) was related to the just preceding positive deviation in a ratio of about 1 to 5. Second, in the enclosed card work (7) reported in 1938, there was found a negative deviation for one experimental condition (correspondence tests with long delay in checking) that was related to the positive deviation of the main series (immediately checked) in the ratio of between 1 to 5 and 1 to 6. Third, in Table IV it may be seen that the correspondence precognition tests with long delay in checking gave, as those just referred to, a negative deviation; in this case the ratio to the comparable laboratory tests is about 1 to 3. (See Appendix.) However, in none of these instances were the conditions giving both trends of deviation comparable and it is somewhat hazardous to speculate upon the equivalent of the factors producing the deviations. Certainly the low-aim studies are the best for comparison, and the deliberately cultured negative deviation in the Adult Division ought to stand a close second.

If it be justifiable to assume for the moment that the most of (and the most comparable of) the experiments dealing with high-aim tests and low-aim tests (including, on the hypothesis of bad aim, the Adult Division) produced approximately similar positive and negative deviations, there is clearly an explanation demanded. Either it must be assumed on the 4 to 1 expectation that subjects trying for low-aim (or with the reversal of aim under the

circumstances of the Adult Division) possess four times the ESP capacity involved in the production of the positive deviation; either that or else there is some compensating factor or circumstance which is at present unknown. It is because of the possibility of experimentally discovering such a factor, if it exists, that the following hypothesis is raised for eventual testing.

If the subject's act in the ESP card tests be assumed to consist of a simple unitary determination of likeness of symbols for high-aim or unlikeness for motivated negative deviation (as Foster's Diametric Function Hypothesis would imply--see article in the preceding issue) there would have to be, to satisfy the 4 to 1 ratio, four times as many successful determinations of unlikeness in motivated negative deviation series (to achieve the same deviation) as there would be successful determinations of likeness in the high-aim. Accordingly, to compensate for the 4 to 1 ratio and achieve the approximately equal deviations actually obtained, there must be supposed in determining unlikeness of symbols in the motivated negative series some greater facility for missing than for hitting when the probability of missing is greater; in fact, it should be proportional to the relative probabilities of missing and hitting. That this is so is in effect the statement of the hypothesis here submitted. If the choices were limited to two, likeness and unlikeness would be equivalent, but the smaller the p value or the greater the range of choice, the less the (ESP) perceptual acuity that is involved in the intentional determination of unlikeness and, correspondingly, the greater the perceptual acuity required for the determination of likeness in high-aim tests.

Specifically, if the subject is working in a low-aim test with ESP symbols the four possibilities for unlikeness are combined, approximately equal psychologically to the one possibility of likeness--i.e., successfully avoiding likeness requires no further consideration as to which of the four possibilities to take. It reduces the task psychologically to a simple dichotomy of approximately similar values, and accordingly, while four times as many successful unlike decisions are required per unit of deviation resulting, there is demanded proportionately less perceptual acuity (or whatever is responsible for extra-chance performance in the tests); and it might be said the task is four times as easy.

It is obvious that this is a problem for experimental study and it is strictly in that interest that it

is discussed here. The hypothesis just offered has, however, a bearing upon another important question; namely, that of accounting for the apparent fact that decreasing the p-value in an ESP test seems also (everything else being equal) to decrease the number of hits in the series attributable to ESP. For a tabular summary of results obtained with different p-values and for a statement of the corresponding ESP quotients (percentages of hits attributable to ESP alone) the reader is referred to page 271 of *Extra-Sensory Perception After Sixty Years*, although he must not be led to expect there anything but the summary of incidental tests with different p-values. No comprehensive experimental comparison has yet been made. Applying the hypothesis of unequal facility in determining likeness and unlikeness in high-aim and low-aim tests it may be seen that the smaller the p-value the relatively more difficult *it should be* to make hits in high-aim series because of the higher perceptual acuity involved in the likeness vs. unlikeness decision inherent in each trial; conversely, the easier *it should be* to determine unlikeness by ESP in motivated negative series, as much easier, in fact, as the ratio of (1 - p) to p. Actually the table referred to will not prove or disprove this hypothesis. There are too many other factors involved to permit any clear-cut effect in any case. But it appears to offer some support to what is given here as a question to be settled by test, and one that may help toward eventual understanding of the psychology of ESP.

Summary

There is reported above the whole of the completed series of precognition tests using mechanical selection of cards, conducted by the author up to March 17, 1939, with the extensive precautions (e.g., two independent experimenters) that have been in operation since the fall of 1938. These series consist (1) of four series totalling 1108 runs of 25 trials each, called the Adult Division, which were conducted to follow up an hypothetical score-depressant factor suspected of producing the negative deviations found in two preceding series mentioned; and (2) of three series totalling 500 runs, called the Child Division, which were carried out as a control, with the expectation, based on earlier work with the same subjects, of obtaining positive deviations. Both lines of expectation were borne out, though at relatively marginal score averages. The grounds for the separate treatment of this work are given and in addition the totality of related but excluded work is summarized for examination in an appendix.

The results, both as reported and as combined either with all the excluded work done with mechanical selection, or with all the known precognition experiments up to June, 1939 (with or without the previously reported hand shuffling work) are statistically significant. The body of results making up this report proper are analyzed in several ways to evaluate the probability of their occurrence by chance: (1) the binomial critical ratio method; (2) the empirical critical ratio; (3) chi-square frequency test; (4) the critical ratio of the difference; (5) the empirical critical ratio of the difference; (6) the combination of the seven series by the chi-square method; (7) the combination of the two divisions by the same method; and finally (8) the binomial calculation of expected sign frequency of the 28 logical sub-divisions of series. In all of these analyses, the results as a whole are significant. Cross-check controls were made on the five of the seven series in which they were possible, with insignificant measures by all the above methods.[10]

The experimental conditions are assumed to have excluded the possibility of selection, of sensory cues, or of errors producing the results obtained. All the important and relevant hypotheses alternative to ESP that are known to the experimenter were discussed. Others may, of course, appear in the course of critical discussion.

To account, then, for the significant relation found between the experimental performance of the subjects and the results, the ESP hypothesis is necessary and appears sufficient; to account for the fact that the coincident results involved the future, since there is no recognized manner in which either rational inference could have given knowledge of the future events, or purposive efforts have brought about the coincident happenings, it is further indicated that ESP is not limited to present events but is precognitive as well. However, considerable restraint is needed in offering and accepting this conclusion, and more evidence is needed to give the hypothesis full and final establishment. The legitimate assumption at this point is that the evidence has reached a stage at which some new alternative hypotheses are needed if precognitive ESP is not to become an established phenomenon.

10. By the 8th method of analysis, that of evaluating the sign or direction of deviation of subseries, the cross-check gives chance results. In the 20th subseries, exactly 10 gave the opposite sign to that experimentally expected; 9 gave the sign expected; and 1 gave no deviation.

Certain incidental findings and suggestions are dealt with, concerning the effect of rewards, of different methods of selection of cards, of the number of subjects participating at a time, the nature of the score-depressant factor, and the interesting and unexpected similarity in the negative and positive deviations per run in the Adult and Child Divisions. None of these minor findings is offered as conclusive, though it is believed they may be of some use to fellow experimenters in the field.

APPENDIX

An effort has been made in the compilation of Table VI to assemble in brief form a general summary of the essential data of all the known precognition experiments conducted up to June, 1939. There is, to my knowledge, left out only the tentative report by Carington of experiments in predicting the throw of a die. This report was issued in a semi-private journal (1) and since it was based on a different p-value from all the other work which we reported, it does not readily pool with the rest.[11]

As stated in the text, the occasion for the presentation of this summarizing table is to present for critical consideration the complete background from which the material for the present paper is drawn. There is as yet insufficient clarity in scientific circles as to exactly what constitutes proper grounds for the separate evaluation and presentation of a particular block of results; although there does not seem to be any reason to question the basis for the separate presentation of the seven series of experiments considered above, it was considered wisest to allow the reader to draw his own conclusions regarding the propriety of selection and to observe that even if no selection whatever is made, the criteria of significance are met by the total results.

The work involved in the assembling of the contents of the table is further justified by the opportunity

11. Since this report was sent off to the Editors, Carington's more recent reports on "Paranormal Cognition of Drawings" have appeared; one in this *Journal* (June, 1940) and the other in the *Proceedings of the Society for Psychical Research* (June, 1940). In these he concludes that evidence of precognition is furnished by the results obtained from his experiments.

it provides for certain general observations over the approaches to the problem of precognitive ESP. Broad comparisons can be made between general conditions and techniques. The explorer may better judge from the inspection of the mass of data given in the greatly reduced form required for tabulation, the relative fruitfulness of various lines of procedure which he might have under consideration, inasfar as these have been followed by others.

The table consists of two sections. Section A is made up of items mostly representing individual researches. This section is divided into two subdivisions, the first dealing with the laboratory methods and the second with the correspondence work. In each of these subdivisions there are two further subdivisions representing the two general methods in use, the calling down through a pack (precognitive DT, or PDT) and second, precognitive matching (or PM). Under the PDT and PM methods there are in each instance four further sub-divisions as well--these representing different techniques of selecting the cards or objects against which the precognitive calling (PDT) or matching (PM) is directed.

Section B is given to general comparisons, first of the two large groups of laboratory experiments and correspondence experiments. First these are compared in their entirety, and then a similar comparison is made with the hand shuffle work excluded. This is done to permit comparison without the question arising as to the possibility of ESP shuffle accounting for the hand shuffle results. A third comparison is made of the total hand shuffle work with the total mechanical selection results. Under each of these headings the distinction is made between the laboratory experiments and the correspondence experiments. Totals of other shuffling methods are presented as well.

There could well be a lengthy discussion made of the results summarized in the following table. The principal point of interest here is the fact that these results are, taken in their entirety, statistically significant. The grand total of 34,810 runs gives a critical ratio of 3.36. Also the chi square combination method gives a highly significant probability ($P = 10^{-12}$). Other ways of evaluating the results involve the comparison of subdivisions of the results and will follow.

Logically, if the results as a whole are significant, it is necessary to know next about the general experimental safeguards. But while they can be said to be

Table VI

GENERAL SUMMARY OF ALL PRECOGNITION EXPERIMENTS UP TO JUNE '39*

Section A

I Laboratory Experiments

A. PDT Division

1. Hand Shuffle

Experimenter	Runs	Dev.	S.D.	C.R.	χ^2
JBR, SOZ, JGP	482	+ 85	± 43.9	1.94	3.76
SOZ, MHP	654	+ 211	51.2	4.12	16.97
GZ, MHP	59	+ 18	15.4	1.17	1.37
SOZ	98	+ 17	19.8	0.86	.74
JBR	45	+ 20	13.4	1.49	2.22
GZ, SOZ	192	+ 63	27.7	2.27	5.15
SOZ	66	- 6	16.2	0.37	.14
SOZ, JBR, GZ	228	- 12	30.2	0.40	.16
JBR	948	- 59	61.6	0.96	.92
GEB, JLM	35	+ 30	11.8	2.54	6.45
EPG	154	+ 49	24.8	1.98	3.92
JGP	1,146	+ 41	67.7	0.61	.37
JDMacF, LMcC, ES	335	+ 165	36.6	4.51	20.34
	46	- 20	13.6	1.47	2.16
Subtotal	4,488	+ 602	134.0	4.49	64.67

2. Mechanical Selection

Experimenter	Runs	Dev.	S.D.	C.R.	χ^2
EPG	14	+ 4	7.5	0.53	.28
GZ, SOZ	70	- 26	16.7	1.56	2.43
JLW (Duke)	1,560	+ 132	79.0	1.67	2.79
JLW (D.M.)	1,897	+ 66	87.1	0.76	.58
CES, JLW	1,794	+ 168	82.7	2.03	4.12
GNMT	90	+ 89	19.0	4.68	21.90
Subtotal	5,425	+ 433	147.3	2.94	32.10

3. Hand Thrown Dice Cut

Experimenter	Runs	Dev.	S.D.	C.R.	χ^2
CES	177	+ 55	26.6	2.07	4.28
JBR, VS	191	+ 47	27.6	1.70	2.89
Subtotal	368	+ 102	38.4	2.66	7.17

4. Newspaper Cut

Experimenter	Runs	Dev.	S.D.	C.R.	χ^2
JBR, JLW	1,652	+ 103	81.3	1.27	1.61
EPG	100	- 4	20.0	0.20	.04
JBR	129	+ 16	22.7	0.70	.49
JLW, BMS	170	- 45	26.1	1.72	2.96
Subtotal	2,051	+ 70	90.6	0.77	5.10
PDT Division Total	12,332	+1,207	222.1	5.43	109.04

Table VI (cont'd)

B. PM Division

1. Hand Shuffle

Experimenter	Runs	Dev.	S.D.	C.R.	χ^2
JGP, EB	45	+ 12	13.4	0.90	.81
LHG, EPG	154	+ 49	24.8	1.98	3.92
Subtotal	199	+ 61	28.2	2.16	4.73

2. Mechanical Selection

Experimenter	Runs	Dev.	S.D.	C.R.	χ^2
BMC, BMS	813	- 69	57.0	1.21	1.46
JBR (Incomp. Miscl.)	236	+ 8	30.7	0.26	.07
" (Adult)	1,108	- 239**	66.6	3.59	12.89
" (Child)	500	+ 123	44.7	2.75	7.56
LER, JGP	282	+ 57	33.6	1.70	2.89
CES, JGP	1,368	+ 98	74.0	1.32	1.74
Subtotal	4,307	+ 456	131.3	3.47	26.61

3. Hand Thrown Dice Cut

Experimenter	Runs	Dev.	S.D.	C.R.	χ^2
JBR (Adult)	80	- 12	17.9	0.67	0.45
" "	400	- 23**	40.0	0.57	0.32
Subtotal	480	+ 11	43.8	0.25	0.77

4. Miscellaneous Selection

Experimenter	Runs	Dev.	S.D.	C.R.	χ^2
JBR (Music)	995	- 33	63.1	0.52	.27
" (Control)	995	+ 35	63.1	0.55	.30
" (Incomp.)	960	- 52	62.0	0.84	.71
Subtotal	2,950	- 50	108.6	0.46	1.28
PM Division Total	7,936	+ 478	176.2	2.71	33.39
Total Laboratory Method	20,268	+1,685	284.8	5.92	142.43

II Correspondence Experiments

A. PDT Division

1. Hand Shuffle

Experimenter	Runs	Dev.	S.D.	C.R.	χ^2
EPG	1,594	- 13	79.8	0.16	.03
Subtotal	1,594	- 13	79.8	0.16	.03

2. Mechanical Selection

Experimenter	Runs	Dev.	S.D.	C.R.	χ^2
JLW	270	- 30	32.9	0.91	.83
EPG	3,188	+ 49	112.9	0.43	.18
Subtotal	**3,458**	+ 19	117.6	0.16	1.01

Table VI (cont'd)

3. Hand Thrown Dice Cut

Experimenter	Runs	Dev.	S.D.	C.R.	χ^2
JGP	126		22.4	0.00	.00
EPG	1,593	- 89	79.8	1.12	1.23
Subtotal	1,719	- 89	82.9	1.07	1.23

4. Newspaper Cut

JLW	531	- 65	46.1	1.41	1.99
EPG	1,591	- 79	79.8	0.99	0.98
JGP	295	+ 11	34.4	0.32	0.10
MMP	226	- 6	30.1	0.20	0.04
JMacF	4,165	- 146	129.1	1.13	1.28
Subtotal	6,808	- 285	165.0	1.73	4.39
PDT Division Total	13,579	- 368	233.0	1.58	6.66

B. PM Division

1. Mechanical Selection

Experimenter	Runs	Dev.	S.D.	C.R.	χ^2
EPG	963	- 64	62.1	1.03	1.06
Subtotal	963	- 64	62.1	1.03	1.06
PM Division Total	963	- 64	62.1	1.03	1.06
Total Correspondence Method	14,542	- 432	241.1	1.79	7.72
GRAND TOTAL	34,810	+1,253	373.1	3.36	150.15

$P(\chi^2) = 10^{-18}$

Section B .

I Comparison of Laboratory and Correspondence Experiments

	Runs	Dev.	Aver.	S.D.	C.R.
Laboratory	20,268	+1,685	5.083	284.8	5.92
Correspondence	14,542	- 432	4.970	241.1	1.79
Grand Total	34,810	+1,253	5.036	373.1	3.36
Diff. Lab. & Corr.	- - -	- - -	.113	- - -	5.20

II Comparison of Lab. Exps. with Corr. Exps. with Hand Shuffle Work Omitted

A. Laboratory Experiments

	Runs	Dev.	Aver.	S.D.	C.R.
Total	20,268	+1,685	5.083	284.8	5.92
Hand Shuffle	4,687	+ 663	1.141	137.0	4.83
Other Selection Methods	15,581	+1,022	5.066	249.6	4.09

Table VI (cont'd)

B. Correspondence Experiments

	Runs	Dev.	Aver.	S.D.	C.R.
Total	14,542	- 432	4.970	241.1	1.79
Hand Shuffle	1,594	- 13	4.992	79.8	0.16
Other Selection Methods	12,948	- 419	4.968	227.2	1.85
Grand Total Less Hand Shuffling	28,529	+ 603	5.021	337.8	1.79
			[P(Comb.X^2) = 1.5 x 10^{-8}]		
Diff. Lab. & Corr.	- - -	- - -	.098	- - -	4.08

III Comparison between Hand Shuffling and Mechanical Selection Results

A. Hand Shuffle Selection

	Runs	Dev.	Aver.	S.D.	C.R.
Laboratory	4,687	+ 663	5.141	137.0	4.83
Correspondence	1,594	- 13	4.992	79.8	0.16
Total	6,281	+ 650	5.103	158.5	4.10
Other Selection Methods	28,529	+ 603	5.021	337.8	1.79

B. Mechanical Selection

	Runs	Dev.	Aver.	S.D.	C.R.
Laboratory	9,732	+ 889	5.091	197.3	4.51
Correspondence	4,421	- 45	4.989	133.0	0.34
Total	14,153	+ 844	5.060	237.8	3.55
Other Selection Methods	20,657	+ 409	5.020	287.5	1.42
GRAND TOTAL	34,810	+1,253	5.036	373.1	3.36

*This manuscript was prepared January-March, 1940. Delay in publication permits the inclusion, in a supplement to this table, of all precognition test results of the year 1939-40, bringing the totals up to October 1, 1940.

** Sign reversed in pooling because negative deviation was expected experimentally.

SUPPLEMENT TO SECTION A, TABLE VI

(Bringing Summary up to October 1, 1940)

I Laboratory Experiments

A. PDT Division

2. Mechanical Selection

Experimenter	Runs	Dev.	S.D.	C.R.	χ^2
LH, JGP	1,000	- 41	63.2	.65	.42
LH	3,100	+ 122	111.4	1.10	1.21
Old Sub-total	5,425	+ 433	147.3	2.94	32.10
New Sub-total	9,525	+ 514	195.2	2.63	33.73

B. PM Division

2. Mechanical Division

Experimenter	Runs	Dev.	S.D.	C.R.	χ^2
JBR	1,880	+ 53	86.7	.61	.37
Old Sub-total	4,307	+ 456	131.3	3.47	26.61
New Sub-total	6,187	+ 509	157.3	3.24	26.98

3. Hand Thrown Dice Cut

Experimenter	Runs	Dev.	S.D.	C.R.	χ^2
JGP, JLW	2,200	+ 90	93.8	.96	.92
LER, DHC, MS	577	- 2	48.0	.04	.0016
Old Sub-total	480	+ 11	43.8	.25	.77
New Sub-total	3,257	+ 99	114.1	0.87	1.69
New Total Lab. Experiments	29,025	+1,907	340.7	5.60	145.35

II Correspondence Experiments

A. PDT Division

5. Miscellaneous Methods of Shuffling

Experimenter	Runs	Dev.	S.D.	C.R.	χ^2
EPG	2,016	+ 136	89.8	1.51	2.28
New Sub-total	2,016	+ 136	89.8	1.51	2.28
New Total Corr. Experiments	16,558	- 296	257.4	1.15	10.00
NEW GRAND TOTAL	45,583	+1,611	427.0	3.77	155.35

$P(\chi^2) \doteq 10^{-12}$

largely of a high order, they are not all equal to the work of the Adult and Child Divisions presented in this report. The only work that offers technically comparable precautions is the Mechanical Selection work under the PM subdivision of the Laboratory Experiments of Section A. Most other work is, in my judgment, essentially almost equal in safeguards; though this might be disputed.

The comparison of laboratory and correspondence experiments in Section B is of special interest because of its confirmation of an earlier finding of similar character (though with non-precognition tests). It may be noted that the correspondence experiments gave a negative deviation, while the laboratory results were positive. This corresponds to the results obtained in the enclosed card work (7) referred to above. In that work the tests in which the checking of the calls and the informing of the subject as to his scores were accompanied by delay (most of them being conducted by correspondence) gave a significant negative deviation, while the immediately checked (laboratory) experiments gave a significantly positive one. In the present case (precognition) the critical ratio of the difference between the two conditions is, as in the enclosed card work, statistically significant (CR Diff. = 5.20).

In the second sub-division of Section B there is given a similar comparison with only the difference that the hand shuffle data are excluded. The elimination of the hand shuffle data does not make any essential difference in the values obtained. The critical ratio of the difference is still 4.08.

In the third sub-division of Section B a comparison of the hand shuffle results and the mechanical selection results is made. In both instances the critical ratios based upon total results for each method are significant, 4.10 and 3.55 respectively.

Although the results summarized in the table are statistically significant, it is by reason of their large numbers that this is so. That is, the average deviation per run is extremely small; namely, .036. This means that whatever is to be assumed as the extra-chance factor producing the significant total deviation, it must be operating almost at the vanishing point. There are at least three positions that might be taken on this point: the first, that the criteria of significance should be raised for work on such an a priori improbable hypothesis as this--raised to a point that would be well above the

probability values ascribed to these mass results. This will not, however, meet with mathematical approval; and further, the probability of 10^{-12} that is given by the chi square combination method exceeds any figure that could reasonably be named as an advanced criterion. Second, it might be suggested that such an average is all that can be expected of precognition tests, that there is something inherently more difficult in them compared to ordinary ESP. But we find that the average obtained by Pratt and Woodruff (6) in their 2400 runs of clairvoyant matching with safeguarding comparable to that of the Adult and Child Division gave an average deviation per run of +.20, which is somewhat smaller than the deviations obtained by both the Adult and Child Divisions (-.22 and +.25). The next series to follow in my own investigations (using the same adult subjects) after the completion of the Adult and Child Divisions was a clairvoyant matching test; the deviation in this was smaller than it had been for the adult precognitive matching (it changed from .22 to .03). Third, it might be supposed that with many investigators conducting exploratory tests in a little known region of problems that there would be increasing "waste motion" as the problems became more obscure and difficult. In preliminary stages, thousands of trials are conducted with very little ground to warrant them, with only the faintest promise and no assurance of significant results--all in the hope of finding a productive lead. This is a necessity in a pioneer research. It might account fully for the "dilution" of average deviation and is at least a reasonable possibility.

But whatever the explanation, it is more important than ever to eliminate possible experimental error, since it would loom proportionately large because of the small margin; and on the other hand to remember that whatever measurable phenomena occur in nature can presumably occur, under some circumstances, in barest traces.

LITERATURE CITED

1. Carington, W. W. "Preliminary Experiments in Precognitive Guessing," *J. S. P. R.*, 1935,29, 86-104.

2. Foster, A. A. "A Perception Ratio Statistic for ESP Tests," *J. Parapsychol, 1940, 4, 320-324.*

3. _____. "Is ESP Diametric?" *J. Parapsychol, 1940, 4, 325-328.*

4. Greenwood, J. A. "On Sampling Procedures Pertaining to Optional Stopping and Selection." (MMs. in the course of preparation)

5. Pegram, Margaret. "Some Psychological Relations of Extra-Sensory Perception," *J. Parapsychol.*, 1937, 1, 191-205.

6. Pratt, J. G., and Woodruff, J. L. "Size of Stimulus Symbols in Extra-Sensory Perception," *J. Parapsychol.*, 1939, 3, 121-158.

7. Rhine, J. B. "ESP Tests with Enclosed Cards," *J. Parapsychol.*, 1938, 2, 199-216.

8. _____. *Extra-Sensory Perception.* Boston: Bruce Humphries, 1934.

9. _____. "Experiments Bearing on the Precognition Hypothesis: I. Pre-Shuffling Card Calling," *J. Parapsychol.*, 1938, 2, 38-54.

10. Rhine, J. B., Smith, B. M., and Woodruff, J. L. "Experiments Bearing on the Precognition Hypothesis: II. The Role of ESP in the Shuffling of Cards," *J. Parapsychol.*, 1938, 2, 119-131.

11. Rhine, J. B., Pratt, J. G., Stuart, C. E., and Smith, B. M., with Greenwood, J. A. *Extra-Sensory Perception After Sixty Years.* New York: Henry Holt and Co., 1940.

12. Smith, B. M. "Note on a Shuffling Machine," *J. Parapsychol.*, 1938, 2, 231-232.

13. Thouless, R. H. "Dr. Rhine's Recent Experiments on Telepathy and Clairvoyance and a Reconsideration of J. E. Coover's Conclusions on Telepathy," *Proc. S.P.R.*, 1935, 40, 24-37.

14. Tyrrell, G.N.M. "Further Research in Extra-Sensory Perception," *Proc. S.P.R.*, 1936, 44, 99-168.

15. Woodruff, J. L., and George, R. W. "Experiments in Extra-Sensory Perception," *J. Parapsychol.*, 1937, 1, 18-30.

Supplement Containing Literature Appearing Since Preparation of this Report

Carington, W., "Experiments on the Paranormal Cognition of Drawings," Proc. S. P. R., 1940, 46, 34-151.

_____. *J. Parapsychol.*, 1940, 4, 1-129.

Hutchinson, L., "Variations of Time Intervals in Preshuffle Card-calling Test," *J. Parapsychol.*, 1940, 4, 249-270.

CONDITIONS AFFECTING ESP PERFORMANCE

Burke M. Smith and Edmond P. Gibson

In this paper are presented, as direct quotations, the opinions, judgments, and conclusions of experimenters concerning what may inhibit or facilitate performance of subjects in ESP tests. This is done in order that lines of further research may be indicated and initiated in the direction of verifying or refuting the more or less tentative statements that have been made. Progress toward understanding the phenomenon of ESP will result largely from actual experimental tests of these suggestions.

As will be seen, there is great unanimity of opinion concerning certain points, and diversity of opinion concerning others. In making comparisons, however, it must be kept in mind that the descriptive terms employed by different authors are not always adequate to ensure the reader of sufficient identity of situations to justify the superficially obvious conclusions.

The main body of this paper is composed of quotations from research articles in ESP in which the authors have stated what to them seem to be pertinent and decisive factors. The material is arranged in a classification into which the items seemed most naturally to fall. There is, of necessity, a certain arbitrariness about the choice of topics under which to arrange the quotations. If the terminology of a particular system of psychology were to be used, there would be great difficulty in determining the relative appropriateness of the various systems and the applicability of the quotations. It was decided, therefore, largely to disregard technical terminology and to group the quotations under headings that are of general current usage, which is the language of the quotations themselves. Considerable latitude is allowed regarding consistency, because the primary objective has been to make this material available without forcing it into any particular systematic framework. There are, indeed a number of overlapping statements; these are indicated by cross references. Although no claim is made to an exhaustive search of all possible sources, it is felt that a reasonably comprehensive survey of the literature of the past sixty years (1880-1939) has been made.

There is given in Table I a list of the conditions which have been suggested by experimenters as important for ESP performance, ranked as to frequency of mention. The figures in the right-hand column show the actual number of quotations bearing upon each condition or group of conditions.

The quotations are alphabetically arranged under the names of the authors of the original reports in which they occurred. The sources are given in the bibliography. In case more than one article by the same author, or joint authors, have been quoted, the exact references are indicated by letters of the alphabet.

TABLE I

A SUMMARY OF QUOTATIONS ON CONDITIONS AFFECTING ESP PERFORMANCE

No.	Conditions	No. of Quotations
I.	Freedom, Pleasure, Excitement, and Interest	24
II.	Disposition, Mood, Personality (Including Health)	23
III.	Variation and Decline of ESP Ability	22
IV.	Abstraction, Relaxation, and Concentration	19
V.	Distraction	16
VI.	Witnesses	12
VII.	Feeling of Success	12
VIII.	Distribution of ESP Ability	9
IX.	Effect of General Fatigue on ESP	8
X.	Role of Agent	7
XI.	Hypnotism and Suggestion	7
XII.	Condition of Stimulus Material	7
XIII.	Distance, Time, and Barriers (Physical)	6
XIV.	Novelty of Materials and Situations	6
XV.	Drugs	6
XVI.	Training (Learning)	5
XVII.	Role of Experimenter	5
XVIII.	Volition	3
XIX.	Observers	3
XX.	Rhythm	3
XXI.	"Action	2
XXII.	Blindness	2

I. FREEDOM, PLEASURE, EXCITEMENT, AND INTEREST

1. Barrett (a) ....but of the favourable effect of freedom from constraint and of a spice of pleasurable excitement, we can speak with entire assurance.

2. Brugmans I believe that the best results with the two-room situation are a consequence of less inhibition.

3. Creery ....and when the children were in good humour, and excited by the wonderful nature of their successful guessing, they very seldom made a mistake.

4. Dessoir ....the agent must be in sympathy with him (the subject), and must himself have the knack of conducting the experiments easily and pleasantly.

5. Estabrooks A fundamental point is the ability to get in quick and friendly rapport with the subjects and put them at their ease as to the experiment, a subject who is nervous or hypercritical or covertly hostile is worse than no subject at all.

6. Gibson No. 4 brought a more intense interest into his experimentation as time went on and he was inclined to be either enthusiastic or discouraged by the tendencies of his scoring.

7. Guthrie ....and as the novelty and vivacity of our seances has departed there is not the same geniality and freshness as at the outset. The thing has become monotonous, whereas it was formerly a succession of surprises.

8. Lodge I asked Miss E. what she felt when impressions were coming freely, and she said she felt a sort of influence or thrill.

9. MacFarland ....one conclusion that is evident is that the mere desire to succeed as an ESP experimenter is not enough.

10. Price and Pegram — The atmosphere during the whole experiment was kept as informal as possible. The subjects viewed the experiment as a game in which each wished to do as well as he could.

11. Rhine, J.B. (a) — ESP may run consistently below chance expectation if there is unconscious (and, of course, if there be conscious) negative tendency of sufficient strength.

12. " — Encouragement is usually helpful. Light humors and moods are the best in which to work at ESP.

13. " — The "curves of operation" found are probably motivational in origin. There is evidence that interest, effort and attention vary, and cause results to vary; there is no clear evidence yet that ESP ability *per se* varies.

14. (b) — One of the most striking similarities between extra-sensory perception and other difficult mental processes is the need in the case of both for confidence.

15. Rhine, L.E. — ....the evident importance for success of psychological setting and "atmosphere" leads one to suspect that a failure may be relative, and may simply be a case of difficulty in finding the most favorable set of conditions. This point, however, is intended as a suggestion and not as an argument.

16. " — In general the children's highest scoring occurred when the conditions approximated most nearly those which obtained in their spontaneous play.

17. Richet — If they seek for lucidity they will not find it..... The answers should come spontaneously and involuntarily....

18. Schmoll & Mabire — It is also indispensable that those who take part in the experiments should not be predisposed toward any plan or theory; that they should interpret the results without bias, but also without exaggerated scepticism; that they should be animated by a sincere desire to discover truth, whatever it may be.....

19. Sidgwick and Johnson — Experiments of this kind (distance) were generally much more wearisome than those at closer quarters, especially as most of them were timed by watches.... and this....entailed a good deal of tiresome waiting on the part of the percipient and of tediously prolonged thinking of the number on the part of the agent.

20. Tyrrell (b) — Experience shows that the occurrence of these brief periods of dissociation (with high scoring) depends upon a psychological background which is largely unknown; but that it is influenced by (i) ideas or suggestions which have been accepted by the subject's subliminal and have taken root in it....(ii) By a general emotional state including interest, enthusiasm, and happiness on the one hand or disappointment, depression and boredom on the other. (iii) By the subject's state of health, which must be good if the right psychological state is to supervene .... That the external conditions of the experiments should include: (i) The possibility of working at a high rate of speed for two reasons, (a) to allow the paranormal impulse to run the gauntlet of conscious associations .... (b) to avoid tedium. (ii) The possibility of dealing rapidly with large numbers of trials in a way which makes the minimum demand on the conscious attention of the subject..... That it is necessary for the investigator to deal with every subject individually and not to attempt to treat subjects *en masse*.

21. Woodruff and George — ....there is some evidence that at least certain subjects do not respond directly to small rewards in ESP tests.

22. " — ....the ability to perform under different conditions varies widely and in different directions among individuals.

23. Van Loon and Weinberg — Emotional processes of consciousness are more easily transmitted than others for those sensations that make the strongest impression on the agent are most easily transferred.

24. Verrall, Mrs. A.W. — The extreme difficulty in keeping up any interest in this "very dull game" may account for the failure of the experiments .....she was keenly anxious to be successful and correspondingly depressed at failure.

II. DISPOSITION, MOOD, PERSONALITY (INCLUDING HEALTH).

25. Barrett (a) — A still more interesting and important question concerns such conditions of success and failure as may lie in the circumstances, disposition, general capacity, and mood of the subject, including such points as consanguinity and familiarity with members of the circle, and also in the temper and manner of the latter.

26. (c) — In fact, when we made the experiments as amusing as possible, we had the greatest success....

27. Bond — No reliable correlation between scoring ability and intelligence rating was found.

28. Creery — (See Item 3)

29. Dessoir — The percipient must be in a calm and contented frame of mind;....

30. Drake — While the cause of this general decline in scoring can be only speculated upon, it is interesting to note, in the light of the observations by Riess, concerning a hyperthyroid subject, that the subject's thyroxin dosage had been doubled a short time before.

31. Estabrooks — The very best type of subject is the one who regards both you and the experiment as a nuisance only to be endured for the sake of gratifying curiosity or for the sake of certain other inducements.

32. Gibson — Following a period of illness her work dropped to chance. It quickly rose again.... and continued at a significant rate....

33. Gurney, Myers, and Podmore	Questions of mood, of goodwill, of familiarity, may hold the same place in psychical investigation as questions of temperature in a physical laboratory.
34. Osty	The best output is obtained when the percipient is morally quiet, in the best possible health, and if the experiment is confined to those times when the subject feels most disposed to work.
35. "	.... the weather, insufficient sleep, some unpleasant news, vexation, slight digestive difficulty--any of these little things may produce psychic trouble sufficient to vary the metagnomic faculty from day to day quite disproportionately to the cause.
36. Rhine, J.B. (a)	.... the definite volitional control over ESP shown by all percipients, the large role of effort and voluntary attention apparent in all .... tend .... to make ESP a part of the natural organization of the species. ESP is directed by the conation of the percipient, and integrates naturally with the other cognitive and with the affective processes of the percipient's mind.
37. "	(See Item 12)
38. Richet	Each subject has his own personal psychology. If we desired to apply to all the same experimental mode, failure would very often result.
39. Riess	Her (Miss S.'s) medical record is of some interest in view of the fact that she gave, as excuse for the lapse of time between the two series of trials, the statement that she had had a "nervous breakdown." Her physician .... gave the following report on her condition. He said that she had had a condition of hyperthyroidism of long standing which he was treating and for which he had prescribed treatment during the three month interval between series..... Miss S. seemed to be a young lady of high intelligence and limited interests. She was obviously not well adjusted to her life in the community.

40. Schrenck-Notzing — On the whole, I have convinced myself that persons who are very suggestible, in whom imagination is active, and particularly those of slightly hysterical temperament, are more suitable as percipients than others.

41. Shulman — Therefore,.... it seems permissible to *suggest* that a clue to the reasons for the significant difference between these two groups (manic-depressive depressed and involutional melancholia) as they scored in this study may be found in the difference in degree of effective attention, external interest, inversion of "thinking and feeling," and negativistic reactions.

42. Sidgwick and Johnson — The reason of the failure is difficult to surmise, unless it was the effect of the distance on the imagination of the agent or the percipient...

43. Thaw — If a percipient can be put in the right condition by a word or two, or by a moment's pressure on the forehead, as in the cases here mentioned,....

44. Tischner — .... and then H. was in very good condition; his health had improved and financial difficulties which had been weighing upon him had come to an end.

45. Tyrrell
(b) — (See Item 20)
(c) — I can only say in the most general terms that the main condition seems to me to be that of mental and physical well-being; the faculty goes with health and happiness, and an absence of any conditions tending to self-consciousness..... joie de vivre.

46. Verrall, Mrs. A.W. — (See Item 24)

III. VARIATION AND DECLINE OF ESP ABILITY

47. Barrett (a) — The delicacy of the conditions was illustrated in our own inquiry partly by the inexplicable fluctuations of success and failure affecting the whole household, partly by the wide difference observed in the capacities of particular members of it from day to day.

48. Comm. on Thought Transference II — ....the average of success in the above experiments, though far above what chance alone could have supplied, falls considerably below the level attained in the trials with the children which were described in Part I of our Proceedings: and .... this decline in power equally showed itself whether the remaining members of the Creery family were or were not cognisant of the object to be guessed. The fact seems to be .... that the capacity is gradually leaving them - a fresh illustration of the fleeting character which seems to attach to this and other forms of abnormal sensitiveness.

49. Comm. on T. T. III — It may be noted that the power of these children, collectively or separately, gradually diminished during these months, so that at the end of 1882 they could not do, under the easiest conditions, what they could do under the most stringent in 1881. This gradual decline of power seemed quite independent of the tests applied, and resembled the disappearance of a transitory pathological condition, being the very opposite of what might have been expected from a growing proficiency in code-communication.

50. Drake — (See Item 30)

51. Gibson — In the case of the principle subject ....there is a remarkable evenness in response to the various experimental procedures and a regularity and consistency which is sustained through a long series of experiments running over two years.

52. Guthrie. (b) — I have noticed a falling off in the facility and success of our experiments since our first great results were obtained.

I do not know to what cause to attribute this declension.

53. Jephson — On the evidence of these curves, if we assume that they are genuine, it does look as if the "fatigue" element became less marked when the clairvoyant faculty is stronger, but it is obviously impossible to draw conclusions from so small a number.

. . . . . .

I consider this definite drop from first sets to last sets in the total results----reinforces very strongly the hypothesis put forward, that we are here observing at work, and progressively deteriorating, an unknown faculty of perception.

54. Osty — But besides persistent moral and organic causes which may bring about persistent weakening of metagnomic faculty, this faculty is found to vary from day to day and even from hour to hour....

55. (See Item 34)

56. Pratt (a) — One subject has been discovered who has given extra-chance scores through many thousands of trials, showing, however, a gradual decline to a chance level of scoring.

57. Pratt and Woodruff — (See Item 170)

58. Rhine, J.B. (a) — The loss of ESP ability with long use is the exception rather than the rule. The ability may decline and return. It may also decline with the daily run.

59. (b) — Capacity for ESP declines when interest declines.

60. (See Items 13 and 74)

61. Richet — Cryptesthesia appears without anyone knowing either why or how. It disappears the same way. It is inapprehensible. We do not know its premises. It is impossible to foresee its destiny. It is very deceptive, but it is a fact.

62. (See Item 38)

63. Tyrrell (a) On the other hand, when the sensitive is engaged in the ordinary occupations of life, telepathy seems sometimes to occur spontaneously....

64. Verrall, Mrs. A.W. (See Item 24)

65. Wales I have noticed some tendency to periodicity in the occurrence of Miss Samuels's veridical impressions. Many days have sometimes passed without anything cropping up in the notes that was applicable to me; at other times, two or three striking correspondences have come close together, even in the same set of notes.

66. Warcollier The phenomena of elementary telepathy which we have been able to produce appear to us to be as independent of the human will as are all the other phenomena of nature.

IV. ABSTRACTION, RELAXATION, AND CONCENTRATION

67. Barrett (a) The common notion that simplicity, and even comparative blankness of mind, are important conditions, seems somewhat doubtfully borne out by our experience.

68. (c) A passive condition of both transmitter and receiver in telepathy seems essential, so far at least as their consciousness is concerned.

69. Brugmans Our subject says that it is necessary that he empty his consciousness (that is his own expression), that it be empty.

70. Despard-Campbell Further, the percipient found she could do better by removing her glasses and closing her eyes, as she was not then distracted by outside objects.

71. Dessoir The percipient .... must endeavor above everything not to strain expectation

in looking for the emergence of the image, but simply to wait quietly. He should empty his brain .... and gaze with closed eyes into a deep darkness.

72. Osty — (See Item 34)

73. Rhine, J.B. (a) — Good abstraction is required for success in most percipients, along with effort and attention to the task; in a word, "concentration."

" — New changes in procedure may disturb this for a time.

Conflict of purposes (and, of course, emotions) spoils concentration.

Dissociation lowers, along with other functions, the capacity to concentrate.

74. " — Loss of ESP ability by an occasional subject may be due to incapacitation for the necessary abstraction under the conditions, rather than a loss of the ESP function itself.

75. (b) — .... our chief problem is to get the subject to let go of his natural inhibitions, his mental habits that are so bound up with rational and sensory processes.

76. " — (See Items 12, 13)

77. Schmoll & Mabire — We would name three conditions as essential to satisfactory experiments: - (1) Absolute silence. (2) On the part of the agents: concentration of thoughts and looks on the objects to be guessed, and firm determination that it shall impress itself upon the percipient's mental vision. (3) On the part of the percipient: complete covering of the eyes, but on no account pressure upon them; perfect calm and an easy position.

78. Sinclair — The next step, after having turned off the light and closed your eyes and relaxed mind and body.....Now, completely relaxed, hold your mind a blank again. Hold it so for a few moments, then give the mental order to the unconscious mind

to tell you what is on the paper you hold in your hand.

79. Thomas .... that is to say, make their minds as much a blank as possible and guess whatever comes into their heads after the signal has been given.

80. Tyrrell (c) G.J. is quite definitely aware,.... of almost losing consciousness of her surroundings. She says that a peculiar, and rather exalted feeling comes over her, making her feel that it is almost impossible to fail, and so long as this lasts, the successes follow one another in an almost unbroken chain; but it is never maintained for more than a few seconds at a time. One sees here a kind of mental dissociation....which lets the extra-sensory faculty overcome its customary inhibitions for a moment or two.

81. (See Item 20)

82. Usher and Burt His method of procedure was as follows:....he....sat down at a table endeavoring to make his mind a blank.

83. Warcollier One must create in oneself a void of thought, keeping the attention solely upon the one idea of visualization.

84. Woodruff and George It seems safe to say that with the subjects here used, conscious attention is important, though not essential, in the efficient operation of this principle.

V. DISTRACTION

85. Bender The presence of an observer materially reduced the number of hits.

86. Despard (See Item 70)

87. Dessoir The preliminary conditions of a successful sitting appear to me to be:- A very quiet locality, with plenty of fresh air and a moderate temperature....

88. Gibson — During this experiment....he.... seemed (to be in) a semi-dissociated state.

89. Gurney et al — (See Item 33)

90. Guthrie — (See Item 7)

91. Lodge — ....but they should be tried soberly and quietly, like any other experiment. A public platform is a most unsuitable place; and nothing tried before

92. Osty — (See Item 35)

93. Rhine, J.B. (a) — There are many facts supporting a view that ESP is easily encroached upon by rational and sensory processes, and that the delicate balance required for good ESP may be concerned chiefly with the maintainance of the field of attention free from these encroachments.

94. (b) — There is apparently no surer way of bringing down ESP scores than by causing nervous dissociation. Whether this dissociation is the result of narcotic drugs, extreme fatigue, or sleepiness does not matter.

95. — (See Items 73, 74)

96. Schmöll & Mabire — Eugenie P: "I do my best to banish any thought that might distract me, and I watch for the appearance. It is not always possible for me quite to renounce thought and volition."

97. — Thus the agents have....to guard against extraneous thoughts which would distract them from their mental labour; while the percipient on his side, cannot sufficiently guard against all cerebral activity; he must carefully deny all free range to his ideas and imagination, and concentrate his endeavours on seizing the impressions or completing the rudiments of those which present themselves to his mental vision.

98. Schrenck-Notzing — The psychical process is comparable to artistic activity. Under the dissecting knife of scientific analysis no artistic creation would be possible: it needs the undisturbed quiet of the studio.

99. Sidgwick & Johnson (See Items 19 & 42)

100. Tischner (See Item 44)

VI. WITNESSES

101. Barrett (See Item 25)

102. Bender The presence of an observer materially reduced the number of hits.

103. Creery Our worst experiments before strangers have invariably been when the company was dull and undemonstrative....

104. Dessoir Only persons ought to take part in the experiments whose presence is agreeable to the percipient, and who he knows will not be disturbed or annoyed by occasional failure.

105. Guthrie (a) It would appear that any mental disturbance on the part of the operators or on the part of the "subjects," due to anxiety to succeed, or to the novelty of the *entourage* of persons or things, very much interferes with the success of experiments.

106. " .... the previous failure having been probably due to the mind of the agent occupying itself with the presence of the visitor, instead of being directed exclusively upon the object to be described.

107. Lodge Usually the presence of a stranger spoils the phenomena, though in some cases a stranger has proved a good agent straight off.

108. Osty It is not even necessary that any antagonism should be manifested. The antipathy of a percipient for some other person is sometimes intuitive and reacts on the psychic faculty. A few more séances generally allow of this repulsion being overcome.

109. Rhine (See Item 73)

110. Sidgwick — The least disturbance of our customary method, change of time or place, presence of strangers, controversy, and expecially noise, is apt to make things go wrong.

111. Tyrrell (c) — But some of the conditions are psychological, very delicate, and largely unknown. Much depends on atmosphere; and, after all, extra-sensory perception is not the only thing which is dependent on atmosphere.

112. Gurney, Myers & Podmore — It must be remembered that we were dealing, not with chemical substances, but with youthful minds, liable to be reduced to confusion by anything in the demeanor of visitors which inspired distaste or alarm....

VII. FEELING OF SUCCESS

113. Adams — ....it would seem that there is no high correlation between subjective certainty and success in attempted ESP.

114. Creery — (See Item 3)

115. Gurney, Myers & Podmore — It is almost inevitable that a percipient who has aroused interest by a marked success on several occasions, should feel in a way *responsible* for further results; and yet any real pre-occupation with such an idea seems likely to be fatal.

116. Jephson — I should expect my group of subjects of being more clairvoyant than the normal ....likely to have had previous experience of clairvoyance and so possibly a bias towards success.

117. Lodge — With regard to the feelings of the percipients when receiving an impression, they seem to have some sort of consciousness of the action of other minds on them; and once or twice, when not so conscious, have complained that there seemed to be "no power" or any thing acting, and that they not only received no impression, but did not feel as if they were going to.

Sometimes they seem quite certain that they are right. Occasionally Miss E. has been pretty confident and yet quite wrong.

118. Richet — Let us not be disheartened, but patient and persevering if the experiments are negative at the start.

119. Riess — She said that she did not believe that there was any difference in scores although she was much less curious about the scores than in the former series.

120. Schmoll and Mabire — ....and resolved to throw into the work the energy that does not recoil in the fact of difficulty,.... and the patience that is not discouraged by a failure, or a whole series of failures.

121. Sinclair — "When the true visions came, there usually came with them a 'something' which I call a 'hunch'."

122. Tyrrell — (See Item 80)

123. Verrall, Mrs. A.W. — (See Item 24)

VIII. DISTRIBUTION OF ESP ABILITY

124. Adams — Published estimates that one person in five has ESP ability seem far too high in the light of these experiments. Here seventy subjects were tested, none of whom was able to demonstrate the ability in question.

125. Comm. on T. T. II — ....we find the capacity present, to a degree admitting of valuable experiment, in a far larger number of persons than we had at first supposed.

126. Creery — I may say that this faculty is not by any means confined to members of one family; it is much more general than we imagine.

127. Guthrie (b) — I have been, I need hardly say, very anxious to find additional cases of

percipiency, and I have induced quite a number of people to make trials in this direction, but hitherto without any satisfactory result.

128. Kubis and Rouke — These physical differences (in twins) would warrant further investigation of the relations of temperament and intelligence to extra-sensory powers.

129. Price and Pegram — Forty-four per cent of the (66) subjects gave individually significant scores.

130. Rhine, J.B. (a) — It seems favorably suggested, at least, that ESP may be heritable or perhaps its more common inhibiting factors may be.

131. Rhine, L.E. — The occurrence of ESP is again demonstrated, in this case....with 17 children aged from three to fifteen.

132. Schrenck-Notzing — The power does not appear to be so rare as is believed; the difficulty lies rather in seizing the moment when such persons are in the right mood.

IX. EFFECT OF GENERAL FATIGUE ON ESP

133. Estabrooks — Using this particular technique it would appear that the faculty which underlies it is extremely susceptible to fatigue, that it works best with a minimum of conscious interference and that any strain or excitement on the part of the percipient is a pretty certain indication of failure.

134. Jephson — (See Item 53)

135. Kubis and Rouke — The factor of fatigue must be avoided, since it seems detrimental to telepathic phenomena.

136. Lodge. — The percipients complain of no fatigue as induced by the experiments, and I have no reason to suppose that any harm is done them. The agent, on the other hand, if very energetic, is liable to contract a headache;....

137. Osty (See Item 34, 35)

138. Pegram The rate of scoring was found to be independent of the number of runs performed each period.....and.... the very last runs of the day had the highest average scores.

139. Rhine,J.B. (a) ESP is not easily fatigued.

140. Tyrrell (a) She had to get herself into a state of great tension, and the number, when it came, was uttered explosively, as it might be by someone who was making a great muscular effort. After giving four or five cards she was usually exhausted....

X. ROLE OF AGENT

141. Creery ....we are all convinced that when mistakes are made the fault rests, for the most part, with the thinkers, rather than with the thought-readers.

142. Dessoir (See Item 4)

143. Guthrie I may here remark that the results of our experience is that success or failure depends as much (if not more) upon the condition of the agent as upon that of the percipient..... When I have not felt equal to the required effort of concentration, I have been unable to repeat the success.

144. MacFarland (See Item 9)

145. Schmoll & Mabire (See Item 77)

146. Warcollier From among our best percipients we chose the person who was to act as agent. She relaxed into the passive state, the usual state for *percipients*.

" Gibotteaus says that one must think of the matter to be transmitted, and then *release it suddenly*. "As agent," he says, "one must discharge one's own mind."

XI. HYPNOTISM AND SUGGESTION

148. Barrett (a) We could find no resemblance between these phenomena and those known as *mesmeric*; inasmuch as a perfectly normal state on the part of the subject seemed our first pre-requisite.

149. (b) When the subject was in the state of trance or profound hypnotism, I noticed that not only sensations but also ideas or emotions occurring in the operator appeared to be reproduced in the subject without the intervention of any sign, or visible or audible communication.

150. Miles and Ramsden There is another point which I do not remember to have heard discussed and that is the influence of suggestion as an aid to the development of psychic faculties. I have found it very effective in my own case.

151. Rhine,J.B. ....we found that we did not need to use the (hypnotic) trance, that it did not help and, of course, made the work much more laborious.

152. Rhine,L.E. It has become increasingly apparent to many ESP workers that suggestion can play an immense part in the outcome of tests of ESP. Unless guarded against with extreme care it can do much to confuse the issue and obscure the true situation.

153. Schrenck-Notzing (See Item 40)

154. Thaw (See Item 43)

XII. CONDITION OF STIMULUS MATERIAL

155. Carpenter and Phalen For a subject who can guess cards with scores beyond chance expectancy, color stimuli are as effective as form stimuli.

156. MacFarland and George It was found that when cards remain unshuffled for a considerable period of time before being used in a test series the scores of subjects will be higher

than when they are shuffled immediately preceding the series. This is offered, however, only as a further problem and in no sense as an established generalization.

157. MacFarland and George — A comparison of scores obtained when normal and distorted material is used, indicates a small difference in favor of normal material largely contributed by two subjects. Even though inconclusive, this would indicate the importance of further attention to situations in which "preference" might be expected to play a more important part.

158. Pratt & Woodruff — Analysis of the data according to size of stimulus symbols used showed that size *per se* did not affect the results of the investigation as a whole.

159. " — (See Item 170)

160. Rhine, L.E. — Within the scope of the experimentation herein reported size variations ranging to proportions of 2,704 to 1 in stimuli did not result in a significant preference in scoring.

161. " — Variations in number of stimuli presented 5 to 1 at a given time did not result in a significant preference.

XIII. DISTANCE, TIME, AND BARRIERS (PHYSICAL)

162. Barrett (a) — Another question would be as to the effect of greater or less distance between the sitters and the guesser, and of obstacles. It will have been seen that, ...the intervention of a door or wall or wall seemed to make no difference.

163. (c) — Another fact which seems to me brought out very clearly in our experiments is that psychical phenomena, whether of telepathy, clairvoyance, or the higher phenomena of spiritualism, are manifestations of, or through, the *transcendental self* of the subject, and are therefore independent of the fundamental units of the physical world-matter, time and space.

164. Brugmans — But in any case, a greater distance does not lessen the phenomenon.

165. Rhine (a) — ....improvement is shown....uniformly with distance, (i.e., when complete failure does not attend it.) Distance discourages sensory attention and would aid abstraction.

166. (c) — ....but the evidence that it (distance) does not influence it (ESP) is nevertheless considerable..... the work already reported seems to show that distance does not limit ESP as it does sensory perception.

167. Sidgwick and Johnson — We were extremely anxious to succeed with agent and percipient at greater distances....but we obtained practically no success.

XIV. NOVELTY OF MATERIALS AND SITUATIONS

168. Estabrooks — It was found that using playing cards, a person very rapidly acquired certain habits which were fatal to the experiment.

169. Guthrie (b) — (See Item 7)

170. Pratt and Woodruff — When the results were considered in relation to the amount of subjects' experience with the different kinds (sizes) of stimulus material, subjects were found to score significantly better with new than with old material. The advantage of working with new material was found to decline after a time as the subjects became more experienced with the new material. Experienced and inexperienced subjects scored equally well with new material. When subjects did not know before or during a run what size of stimuli were used, the advantage in favor of higher scores with new material was not statistically significant. The evidence suggests that there was a period of adjustment during the first runs in the first sessions of using any material, whether new or old.

171. Rhine, J.B. Variation of the procedure, without introducing new changes, helps to keep up the scoring level.

172. " (See Item 73)

173. Verrall, Mrs. A.W. (See Item 24)

XV. DRUGS

174. Brugmans As for the alcohol, we can say that it augments the results considerably.

175. Rhine, J.B. Dissociative drugs, sleepiness and certain illnesses work to lower this integration and self-control; whereas drugs that antagonize dissociative drugs help to recover normal control. And in our results the data show plainly that dissociative factors likewise lower ESP ability, while counteractive factors help to restore it.

176. " The effect of sodium amytal has been rather strikingly destructive to ESP....

177. " ....the effect of caffeine was equally striking. The first treatment brought his average up from 12.8 to 14.7 in 300 trials..... But the best test for caffeine came 5 hours after taking the sodium amytal. Under the influence of the amytal he had dropped to an average of 6.2 in the last 300 trials. One hour after the 5-gr. dose of caffeine, he rose to an average of 9.5 ....

178. (See Item 94)

179. Tanagras It is evident that, in this circumstance, the player, finding himself, thanks to the champagne, in a light "Absence," emitted psychological energy....

XVI. TRAINING (LEARNING)

180. Pratt and Woodruff — The amount of experience of subjects with ESP tests was not, in itself, directly associated with trends in scoring.

(See Item 170)

181. Richet — I believe that all our efforts to intensify this faculty in a sensitive are barren and even sometimes disastrous....

182. Rhine (a) — Improvement in ESP is limited, usually to a short initial period; and this may be purely a matter of learning to abstract and achieve good concentration. There is no good evidence of improvement of the ability. A level seems to be reached and not far exceeded. When it is exceeded for particular runs, it seems to be through special effort.

183. Warcollier — Briefly, the lesson learned from these experiments in telepathy is that no telepathic training is possible for percipients, except for them to develop their imagination to permit the telepathic message to manifest itself.

XVII. ROLE OF EXPERIMENTER

184. MacFarland — But the two experimenters still continued at approximately their previous levels of success, one at chance and the other consistently above.

185. — The research shows only that if differences in personality were responsible for the results, they were at least not dependent on direct contact with subjects during the tests; experimenters and subjects had no such personal contact.

186. Pratt and Price — The writers conclude that failure to find evidence for ESP in card tests may be due to an unfavorable experimenter-subject relationship.

187. Sharp and Clark — It was observed that marked variation in group scoring coincided with changes

in the experimenter's attitude. Individual scoring showed notable changes with different experimenters.

188. Sharp and Clark — During these weeks the experimenter was working, then, with divided attention (His wife was ill.)..... It is interesting to note that each week under "normal" conditions gave a positive deviation; each "abnormal" week, a negative one.

XVIII. VOLITION

189. Pegram — Low-aim showed more consistent deviations from chance than did high-aim.

190. Rhine — (See Item 36)

191. Warcollier — (See Item 66)

XIX. OBSERVERS

192. Barrett (a) — It would be interesting, again, to discover whether numerical increase in the observers increases the effect, and how far the presence of the father....seemed decidedly to increase the percentage of successes.

193. Dessoir — (See Item 87)

194. Rhine — (See Item 73)

XX. RHYTHM

195. Barrett (a) — On one or two occasions it seemed of advantage to obtain vivid simultaneous realisation of the desired word on the part of all the sitters; which is most easily effected if some one slowly and gently claps time, and all mentally summon up the word with the beats.

196. Stuart — It is suggested that the (highly significant) difference in scoring level noted is due to distraction of attention from the primary task of matching when imposed non-tempo rates require effort to keep in time.

197. Tyrrell — Rhythm in working through a series of trials is desirable.

XXI. ACTION

198. Tyrrell — (See Item 20)

199. Verrall, H. de G — The Misses Tipping are both under the impression that more success is likely to be attained when action of some kind is involved in the subject of the experiment, and the results hitherto observed in this case bear out this impression.

XXII. BLINDNESS

200. Price — The average scores of the tests given the blind subjects are only slightly but fairly consistently higher than those obtained from the non-blind.

201. Price and Pegram — No reliable correlation was noted between ESP scoring ability and age, extent of blindness, nor with the time in which the subject became blind.

REFERENCES

Adams, E.T. "A Summary of Some Negative Experiments," *J. Parapsychol.*, II (1938), 232-36.

Barrett, W.F.
(a) "First Report on Thought Reading," *Proc. S.P.R.*, I (1882), 13-34.
(b) "On Some Phenomena Associated with Abnormal Conditions of Mind," *Proc. S.P.R.*, I (1883), 238-44.
(c) "Some Reminiscences of Fifty Years' Psychic Research," *Proc. S.P.R.*, XXXIV (1924), 275-97.

Bender, Hans. (Review by Adams, D.K.) "Bender on Extra-Sensory and Sensory Form Perception," *J. Parapsychol.*, I (1937), 52-62.

Bond, Esther, "General Extra-Sensory Perception with a Group of Fourth and Fifth Grade Retarded Children," *J. Parapsychol.*, I (1937), 114-22

Brugmans, H. I. F. W. "Une Communication sur des Experiences Telepathiques au Laboratoire de Psychologie à Groningue Faites par M. Heymans, Dr. Weinberg, et Dr. H. I. F. W. Brugmans, "*Compte Renau Officiel du Premier Congrès International des Recherches Psychiques* (1922), 396-408.

Carpenter, C. R. and Phalen, Harold R. "An Experiment in Card Guessing," *J. Parapsychol.*, I (1937), 31-43.

Committee on Thought Transference. "Second Report," *Proc. S. P. R.*, I (1882), 70-98.

. "Third Report," *Proc. S. P. R.*, I (1883), 161-216.

Creery, A.M. "Note on Thought Reading," *Proc. S. P. R.*, I (1882), 43-46.

Despard, R. C. and Campbell, C. M. "Experiments in Thought Transference," *J. S. P. R.*, VI (183), 4-9.

Dessoir, Max. "Experiments in Muscle-Reading and Thought-Transference," *Proc. S. P. R.*, IV (1886), 111-126.

Drake, Raleigh M. "An Unusual Case of Extra-Sensory Perception," *J. Parapsychol.*, II (1938), 184-98.

Estabrooks, G. H. "A Contribution to Experimental Telepathy," *Bull. Boston S. P. R.* (1927).

Gibson, E. P. "A Study in Comparative Performance in Several ESP Procedures," *J. Parapsychol.*, I (1937) 264-75.

Gurney, Edmund; Myers, F. W. H.; and Podmore, F., *Phantasms of the Living*. London: Trubner & Co., 2 vols., 1886.

Guthrie, Malcolm.
(a) "An Account of Some Experiments in Thought Transference," *Proc. S. P. R.*, II (1883), 24-42.
(b) "Further Report on Experiments in Thought Transference at Liverpool," *Proc. S. P. R.*, III (1885), 424-52.

Jephson, Ina. "Evidence for Clairvoyance in Card Guessing," *Proc. S. P. R.* XXXVIII (1929), 223-68.

Kubis, J. F., and Rouke, F. L. "An Experimental Investigation of Telepathic Phenomena in Twins," *J. Parapsychol.*, I (1937), 163-71.

Lodge, Oliver J. "An Account of Some Experiments in Thought Transference," *Proc. S. P. R.*, II (1884), 189-200.

MacFarland, J. D. "Discrimination Shown Between Experimenters by Subjects," *J. Parapsychol.*, II (1938), 160-70.
and George, R. W. "Extra-Sensory Perception of Normal and Distorted Symbols," *J. Parapsychol.*, I (1937), 93-101.

Miles, Clarissa, and Ramsden, H. "Experiments in Thought Transference," *Proc. S. P. R.*, XXI (1908), 60-93.

Murray, Gilbert. "Pres. Address," Proc. S. P. R., XXIX (1918), 46-63.
Osty, Eugene. *Supernormal Faculties in Man.* London: Methuen, 1923. 134-35.
Pegram, Margaret H. "Psychological Relations of Extra-Sensory Perception," *J. Parapsychol.*, I (1937), 191-205.
Pratt, J. G.
(a) "Clairvoyant Blind Matching," *J. Parapsychol.*, I (1937), 10-17.
(b) "The Work of C. Hilton Rice," *J. Parapsychol.*, I (1937), 239-59.
, and Price, M. M. "The Experimenter-Subject Relationship in Tests for ESP," *J. Parapsychol.*, II (1938), 84-95.
Pratt, J. G., and Woodruff, J. L. "Size of Stimulus in Extra-Sensory Perception," *J. Parapsychol.*, III (1939), 121-58.
Price, M. M. "A Comparison of Blind and Seeing Subjects in ESP Tests," *J. Parapsychol.*, II (1938), 273-86.
, and Pegram, M. H. "Extra-Sensory Perception Among the Blind," *J. Parapsychol.*, I (1937), 143-55.
Rhine, J. B.
(a) *Extra-Sensory Perception.* Boston: Bruce Humphries, 1934.
(b) *New Frontiers of the Mind.* New York: Farrar and Rinehart, 1937.
(c) "Effect of Distance in ESP Tests," *J. Parapsychol.*, I (1937), 172-184.
Rhine, L. E. "Some Stimulus Variations in Extra-Sensory Perception with Child Subjects," *J. Parapsychol.*, I (1937), 102-113.
Richet, Charles. *Our Sixth Sense.* London: Rider & Co., 1927.
Riess, B. F. "A Case of High Scores in Card Guessing at a Distance," *J. Parapsychol.*, I (1937), 260-63.
Schmoll, A., and Mabire, J. E. "Experiments in Thought Transference," *Proc. S. P. R.*, V (1888), 169-215.
Schrenck-Notzing, Baron von. "Experimental Studies in Thought-Transference," *Proc. S. P. R.*, VII (1891), 9-22.
Sharp, V., and Clark, C. C. "Group Tests for Extra-Sensory Perception," *J. Parapsychol.*, I (1937), 123-42.
Shulman, Robert. "A Study of Card Guessing in Psychotic Subjects," *J. Parapsychol.*, II (1938), 95-106.
Sidgwick, Mrs. Henry. "Report on Further Experiments in Thought-Transference Carried out by Prof. Gilbert Murray," *Proc. S. P. R.*, XXXIV (1924), 212-274.
Sidgwick, Mrs. Henry, and Johnson, Alice. "Experiments in Thought Transference," *Proc. S.P.R.*, VIII (1892), 536-96.

Sinclair, Upton. *Mental Radio.* Pasadena: Upton Sinclair, 1930.

Stuart, C. E. "The Effect of Rate of Movement in Card Matching Tests of Extra-Sensory Perception," *J. Parapsychol.*, II (1938), 171-183.

Tanagras, A. *Le Destin et la Chance.* Société des Recherches Psychiques Hellenique: Athens, 1938.

Thaw, A. Blair. "Some Experiments in Thought Transference," *Proc. S. P. R.*, VIII (1892), 422-35.

Thomas, N. W. *Thought Transference.* London: Alex Moring, Ltd. 1905.

Tischner, Rudolph. *Telepathy and Clairvoyance.* (Trans. by Hutchinson, W. D.) New York: Harcourt, Brace, 1925.

Tyrrell, G. N. M.
(a) "The Case of Miss Nancy Sinclair," *Jour. S. P. R.*, XX (1922), 294-327.
(b) "Further Research in Extra-Sensory Perception," *Proc. S. P. R.*, XLIV (1936), 99-168.
(c) *Science and Psychical Phenomena.* New York: Harpers, 1938.

Usher, F. L., and Burt, F. P. "Thought Transference," *Annals of Psychical Science*, VIII (1909), 561-600

Van Loon, F. H., and Weinberg, A. A. "A Method of Investigation into Thought Transference," *Jour. S. P. R.*, XX (1921), 34-49.

Verrall, Mrs. A. W. "Supernormal Acquisition of Knowledge," *Proc. S. P. R.*, XI (1895), 174-93.

Verrall, H. De G. "Some Recent Experiments in Thought Transference," *Proc. S. P. R.*, XXVII (1915), 415-457.

Wales, Hubert. "Apparent Thought-Transference," *Proc. S. P. R.*, XXXI (1921), 124-217.

Warcollier, René. *Experimental Telepathy.* Boston: Boston S. P. R., 1938.

Woodruff, J. L., and George, R. W. "Experiments in Extra-Sensory Perception," *J. Parapsychol.*, I (1937), 18-30.

Board of Review, Journal of Parapsychology

Memorandum to Dr. Gardner Murphy, Editor

Subject: Review of "Experiments Bearing on the Precognition Hypothesis, III Mechanically Selected Cards," by J. B. Rhine

This manuscript has been reviewed by the Board and the following members have submitted criticisms which are incorporated in this memorandum; E. R. Hilgard, R. R. Willoughby, I. Lorge, J. L. Kennedy, L. Long and S. B. Sells. Specific points are credited to each reviewer; all other statements are the responsibility of the writer.

The criticism covers the writing and organization of the report, the design of the experiment, the statistical method of evaluating the results and the interpretation.

A. Writing and Organization of the Report[1]

The report is poorly written from the standpoint of following the experimenter's trend of thought and keeping track of details of procedure and interpretation.

Dr. Willoughby considers this paper a retrogression on the part of the Duke Laboratory after the highly satisfactory reporting of their experiments by Pratt and Woodruff last year. "The style is back in the phase where you had to hunt through a maze of explanations as to whether it is preferable to believe that the subjects affected the shuffling by remote control or received advance information as to what the shuffle was going to produce, etc. ...to locate such a critical point as whether they were trying to match or not to match the cards."

Dr. Lorge has made a similar criticism. "It is extraordinarily difficult for a person not conversant with the literature to understand either the problem, the

1. Certain of the detailed recommendations of the Board *concerning the writing* have been followed in the revised report. These are omitted from this memorandum, as indicated. The writer believes that the present revision is superior to the original in its presentation. Unfortunately time did not permit consideration of the revision by the Board. (April 7, 1941)

procedure, or the conclusions. The paper is unduly verbose, vague and ambiguous. I recommend that Professor Rhine be advised to rewrite the paper in the form of hypothesis, justification for hypothesis, procedure necessary to test the hypothesis, description of procedure exactly with illustrative material so that a person not having had previous experience in this line of research might understand as he reads this paper, analysis of the data and conclusions, conclusions, however, to be directed to the data from this study first and later integrated with such other relevant literature as the author deems advisable." And again. "Furthermore, data should be incorporated at the point of interest and not included as an appendix. This is an article, not a monograph."

Dr. Lorge made a detailed list of criticisms of the organization of the paper. These have been carefully checked and are endorsed in toto by the writer. (The list originally included has been omitted by the writer since most of the suggestions included have been incorporated in the revision of the report. April 7, 1941. S. B. Sells)

B. Design of the Experiment

The design of the experiment has been criticized by every reviewer. Part of the criticism has been reflected in the preceding statements concerning organization and writing.

Dr. Hilgard points out the general laxity of procedure. "The general arrangements, while admittedly not meeting all the requirements for independent records suitable for later checking, appear to be generally satisfactory. There is still an atmosphere of informality about arrangements, such as the author being called out of the experimental room to answer telephone calls (ms.p.6). Any "later" discrepancies as mentioned on p.11(ms.) should be subject to review through examining original records, not through arbitrarily choosing a value favoring the chance mean."

Dr. Kennedy has reported that he observed no obvious experimental errors, but that he felt that failure to mention one condition of control throughout the experiment was a fault in technique.

Dr. Long did not observe any serious omission in the experimental procedure. However, he criticized the design as follows: "The determination of the number of the correct responses by chance alone in this paper

seems to run into the same difficulty that has occurred before in ESP experiments. If chance is 5 in 25 then it must be assumed that there are 25 independent choices. Actually, this is not the case. With no screen or other sensory precautions it is obvious that the subject worked with knowledge of the numerical distribution of the cards. Certainly it is not difficult to believe that the subject would refrain from placing 6 cards in any one pile. Would it be possible to give a subject a pack of 25 cards and tell him to select only 5 cards? That is, he is to place only one card which is later to be matched with the key card. He can take the first 5 cards or he can run through as many cards as he likes, but only five cards are to be placed out front. This would make the experiment a little longer and cut down on the N, but on the other hand Rhine seems to inflate the N by his procedure. In the suggested situation it seems that chance would definitely be 1 in 4 since the judgments are independent."

Dr. Willoughby has raised a similar point. "We are led back to the point binomial and the square root of npq. It seems to me that the Huntington-Sterne and Kellogg contributions successfully replaced this approximation with the precisely appropriate distribution, at least for the standard situation. The situation now discussed seems to me not to be the standard one. In the standard situation the calls vary in a very large number of ways, in a number of ways limited by the habits of the subject, and the cards vary independently in the same number of ways; it was the achievement of the Huntington-Sterne group and of Kellogg to determine the proportion of each kind of combination that results from all possible associations of these two very large variables. In the present situation the calls vary in something like the same way--except that the habits of persons sorting cards into five groups are pretty certainly different from those of the same persons reporting on the images of these cards that occur to them--but it seems to me that the "cards," i.e., the target series of five to be determined by post-guessing shuffle, vary in only 5.4.3.2.1 or 120 ways. If this is correct, the chances of making a hit would be greatly increased. What we need, of course, is a distribution of the chances of making 9, 1, 2, 3...hits, and on the same axis a distribution of the hits actually made--whatever the correct chance distribution."

One of the most important criticisms of the design of the experiment is what we may call its lack of design. Many different conditions are employed,

different subjects, varying in age and background are tested and the results are pooled. No individual records are presented.

Dr. Lorge has written an important criticism on this point. "May I suggest that as I understand this piece of research, the variables are:

1. Shuffle machines, versus
2. dice machines;
3. reward, versus
4. non-reward;
5. adult, versus
6. child;
7. individual, versus
8. group.

The design of the experimental layout could have been arranged in Latin-Greek squares so that various sources of variation could be evaluated at least tentatively. The author unfortunately through poor design may be making inferences come in favor of one or the other of his variables. I believe that individual variation is a source of considerable ESP effect which is not necessarily extrasensory. I think that contingency or correlation methods are absolutely necessary to settle this issue."

C. Statistical Method and Interpretation

Part of the criticism in this section is contained in the preceding statements.

Dr. Hilgard has raised the following points: first, "Objection may be raised to pooling raw discrepancy scores with sign reversals as in Tables I, II, III, and VI. The fact is clearly indicated in each case, but the reason given, that the result was 'experimentally anticipated' means in operational terms only that it was experimentally found. It was not even anticipated in terms of a prior hypothesis, although even such anticipation would not justify pooling. This objection does not apply to Chi Square treatments, but does apply to summating raw discrepancies from expected mean values. By the same logic, all signs could be made alike, provided only that a consistent sign was found for one stated circumstance of experimentation"; and second, "reiteration is needed for the warning on p. 3 of the appendix that the average deviation for the data from precognition experiments, as pooled with the sign reversals, is only .036, so that 'whatever is to be assumed as the extrachance factor producing the significant total

deviation, it must be operating almost at the vanishing point.'"

Hilgard's first point was noted also by other members of the Board. Dr. Lorge "urgently requests the author not to advise the reader that material is unimportant and then discuss such material or that the logic for an experiment is not relevant and then infer in his conclusions about the necessary logical steps." Dr. Willoughby states "we need give no special attention to the fact that whether the subjects (instructed to match the cards, we eventually find out) succeed or fail in the task makes no great difference, since the results are interpreted as evidence for ESP anyway. We might accept such reasoning if actual distributions were presented and shown to be V-shaped. It is impossible to accept it on the basis only that "it was experimentally anticipated," that is, that Dr. Rhine thought in advance (he says) that it might turn out that way." The same point was made by Dr. Sells.

Conclusion

On the basis of the above criticism it is the opinion of the writer that the paper be referred back to Dr. Rhine for revision. A copy of this memorandum may be forwarded to him for any guidance he may wish to accept ffom it. In the event that it is inexpedient to revise the paper, it is requested that this memorandum be published simultaneously with it.

Saul B. Sells, Chairman

Department of Education
Division of Graduate Studies
Brooklyn College

To the Editors:

I have gone over the criticisms made by the Board of Review on my report published in this issue, and I have tried to do everything reasonably possible to conform to what (even an individual member of) the Board thought it worthwhile to suggest. This has not meant, of course, that I have been able to see the justice of all the criticisms or the need for the changes made. Since the report covers a somewhat complicated research, naturally some patience is called for on the part of the reader if the organization is to be followed. Further, the demands of the great number of counter-hypotheses with which the ESP experimenter is now contronted requires emphasis upon *completeness* of reporting to a degree that suggests *verbosity* to the reader who does not have the various counter-hypotheses in mind.

These considerations may not excuse, though they may qualify, the impression of wretched reporting which some members of the Board of Review seem to have got from my paper. Having been at some pains to amend the manuscript wherever possible to meet the criticisms, I trust I am not letting myself off too easily to conclude that after all if the report is understandable at all, the most important question is: Is the conclusion warranted? For this question the second and third sections of the report of the Board of Review are necessarily of consideration.

The sections on "Design of the Experiment" and "Statistical Method and Interpretation" contain some criticisms which I am unable to comply with, in spite of my appreciation for the time and consideration which the members of the Board have given to the matter. Accordingly I think it would be best to refer the issues to the judgment of the reader by offering brief comments on the specific criticisms for publication with this letter.

Yours very truly,

J. B. Rhine

Remarks on the Criticisms Summarized under "Design of the Experiment"

Under the heading "Design of the Experiment," Dr. Sells apparently includes both (a) the question of whether the experimental procedure is adequate for the testing of the hypothesis and (b) the question of whether the experiment lends itself satisfactorily to statistical analysis.

On the former question of experimental adequacy, there is no serious criticism given. In view of the conclusions of the report, *this is of the utmost importance.* As a matter of fact, Dr. Long definitely states that he did *not* observe any serious omission in the experimental procedure. Of the two who did, one of them, Dr. Kennedy, made only a mild criticism. Dr. Hilgard's criticisms concern matters of nicety, since he himself uses the expression "generally satisfactory" to apply to the general experimental procedure. I have reviewed the situation because I am at a loss to understand Dr. Sells' interpreting Dr. Hilgard as implying "general laxity of procedure."

Dr. Lorge's criticism of the design of the experiment (his impression of the *lack* of design, as Dr. Sells puts it) is dealt with adequately, I think, in the revision of the manuscript--particularly by the inclusion of Figure 3 representing graphically the design followed. From this it will be seen that the experiment is complex but by no means confused.

However, in the section on "Design" there are several criticisms made which are statistical in character. Dr. Lorge suggests the use of the Latin square method, but I can discover no advantage which it would have in this research. Every square in the experimental design is properly evaluated--if need be, against a theoretical chance background; and, for intersquare relations, comparison and combination of critical ratios are feasible. For that matter the data are amenable to an Analysis of Variance treatment if that seems desirable.[1]

1. Dr. Greenwood has applied the Analysis of Variance technique to the experimental results under discussion. He contemplates submitting a report on this treatment to the editors of the Journal with a view to acquainting ESP workers in general with the use of this method. This, however, is not the Latin square treatment which Dr. Lorge specifically suggested but is an outgrowth of his suggestion.

Likewise I can see no ground nor discover any mathematical authority for Dr. Lorge's emphatic statement that contingency or correlation methods are absolutely necessary to settle the issue of this report. These methods may be, and on occasion have been, used; but not from absolute necessity.

The old question of the applicability of the matching hypothesis is brought up again specifically by Dr. Willoughby and indirectly by Dr. Long (in his question of the independence of the twenty-five choices). I think it is fair, as it certainly is most economical, to refer this question back to where it was settled in the first place--by mathematicians of probability. For Dr. Greenwood's summary, see Chapter II, "Extra-Sensory Perception After Sixty Years."

The two remaining questions of statistical character raised under the heading of "Design" both lead to the same solution; namely, to the cross-check series. These are: first, Dr. Long's question whether in matching tests the number of cards matched represents the real number of trials or whether there is a padding effect; and second, Dr. Willoughby's question as to whether the matching (sorting) techniques do not require a special empirical distribution of run scores in order to determine the type of distribution for this set of data. The answer is that the empirical control distributions (both for matching and for calling methods) available in the now numerous cross-check series are applicable to such issues as those raised. Dr. Long's problem is essentially touched upon in the section entitled "Stacking Effect." (See Table II of the report for the cross-checks; for similar material see Table II, "Extra-Sensory Perception After Sixty Years;" also Miss Hutchinson's report in the preceding issue of this Journal.)

Remarks on Criticisms on "Statistical Method and Interpretation"

Under this heading the criticisms center upon the question of the statement of objectives given for the Adult Division. I stated simply that having got negative deviations in two comparable introductory experiments, I was seeking for a negative deviation throughout the four series of the Adult Division.

Now I doubt seriously if any one of the critics supposes that I went through the six negative series completely blind to consequence. Nor can he well suppose that I should be quite so unintelligent as to continue

so long steadily rolling up a large negative deviation while seeking for a positive one. The only reasonable question is as to the point where the induction that such a procedure would lead to a negative deviation occurred, and where a negative deviation was essentially predicted for the next series. My statement of intention, as well as the objective indications, show that this was at the point following Introductory Series A and B.

But Dr. Hilgard pointed out what Drs. Sells and Willoughby apparently did not recognize; namely that the sign of deviation is not of concern in the chi square treatment; and it is obvious that the conclusions may rest upon chi square treatments alone.

Moreover, regardless of statements of intention the two divisions may be treated separately by the critical ratio method. There is no question of the legitimacy of this and since each division is independently significant by this method, there is no need to dispute the propriety of pooling the deviations with sign reversal. Accordingly, I have deleted the data from pooling with sign reversals from Tables I, II, and III. (I have not done so for the Appendix Table VI since this table is given *for the purpose* of pooling of all totals and is intended to be complete. The only alternative here to pooling with sign reversal would be pooling without sign reversal, but this would be to assume the *direct opposite* of the actual experimental objective for the Adult Division. Again, however, the statistical significance of the total results does not rest upon the point discussed.)

The use of significant negative deviations to support the ESP hypothesis appears to disturb Drs. Willoughby and Sells. This matter is discussed at length in the later sections of the report under discussion (to which no reference is made by the Board). In addition their attention might be directed to familiar examples of *successful avoidance responses* as exhibited, for instance, in hypnotic blindness demonstrations in which the subject successfully avoids contact with the objects he is supposed not to be able to see. Obviously a card test under such conditions would show, not chance coincidence, but a negative deviation from chance. The results would presumably be construed as evidence of perception--of a kind.

To the Editors:

I should appreciate the courtesy of space in the Journal in order to answer a charge brought against me by Dr. Willy Feller in his article: Statistical Aspects of ESP, which appeared in the December issue.

He states, on page 287, that the validity of the optional stopping effect in ESP experimentation "was again and again vigorously denied," and then, apparently as an instance of such a "denial," quotes in a footnote a paragraph from my paper at the ESP symposium in 1938. As this quotation contains no reference to the effect usually described by the term "optional stopping," I feel that he owes it to me and to his readers to explain how he is able to so construe it. His startling statement as to the probability of attaining a given success ratio also deserves to be amplified.

The optional stopping effect generally discussed in connection with ESP data (and the only one considered by Dr. Feller in his article) is concerned with the possibility of obtaining a high value of the so-called *"critical ratio"* by choosing a favorable stopping point. On the other hand, the paragraph from my paper deals with the probability of obtaining a given success ratio in a prolonged series: a situation which might be held to involve one form of "optional stopping," but is governed by an entirely different set of considerations.

The quotation in question reads as follows:

> "The statement is occasionally made that by continuing long enough one is certain eventually to obtain a fairly high average total score. This is clearly at variance with the recognized laws of probability, but in order to meet the objection more specifically I have worked out a formula which gives, with a high degree of accuracy, the probability that a given score will ever be reached in a series of infinite length....For example, the probability in an infinite series of calls of ever securing, at any time after the first 10,000 calls, a total average score of 5½ correct guesses per 25 can be shown to be less than one in a million."

Immediately following the above quotation, Dr. Feller says:

> "In reality this probability is 1 (certainty) and one is certain to get, infinitely often, an arbitrarily large C. R."

That the unambiguous wording "total average score of 5½ correct guesses per 25" should be misunderstood by anyone well read in the literature of ESP is most unlikely. Yet it is equally incomprehensible that it should be necessary to point out to a mathematician of Dr. Feller's standing in the field of probability the absurdity of the statement that there is a unit probability of obtaining a success ratio greater than the probability of success on a single trial. That this cannot be true is at once apparent if we compute the C. R. necessary to produce an average score of 5½ per 25. This is $1/50 + 2/5\sqrt{n}$ which equals $\sqrt{n}/20$. Now, as Dr. Feller himself points out in his footnote 32, the order of magnitude of the greatest observed C. R. in a series of n trials is expected to be $\sqrt{\log \log n}$, while the C. R. required for a success ratio of 11/50 has just been shown to be of the order $\sqrt{n}$, which increases much more rapidly than $\sqrt{\log \log n}$. The derivation of the precise figure of "less than one in a million" is too technical to reproduce here, but follows easily from the usual formulae for the probability of the ruin of one of two gamblers betting on the outcome of a series of games, where their probabilities of winning are constant from game to game (see, for example, J. V. Uspensky, *Introduction to Mathematical Probability*, p. 146).

It is to be hoped that Dr. Feller will explain in detail just what he meant by this footnote.

T. N. E. Greville