The Journal *of*
Parapsychology

Volume 2 DECEMBER, 1938 Number 4

WILLIAM McDOUGALL (1871-1938)

Professor William McDougall, head of the Department of Psychology, Duke University, and one of the editors of this JOURNAL, died on November 28, 1938 at the age of sixty-seven years. A brief review of what he has done for parapsychology (and psychical research in general) will make clear that his death is an incalculable loss to this field of study.

At the same time, to confine such a review merely to parapsychological matters is to give but a very slender segment out of the rich and extensive life of contribution which Professor McDougall has lived. Parapsychology, although it interested him greatly, occupied a relatively secondary place among his many interests. He once referred to it as one of his two hobbies. In spite of this, he became one of its most outstanding figures.

Professor McDougall's interest in the problems of parapsychology dated back to his boyhood when, as a skeptic, he learned to regard lightly the various spiritualistic manifestations popular among the mediums of the day. Later on, however, as a student at Cambridge University, he acquired a more serious interest in the problems that were being dealt with by the Society for Psychical Research, and he became a member of the Society in 1901. During the first quarter of the century, he wrote a number of reviews for the Society's publications and otherwise evinced active interest in its problems and development. In 1920, he was elected to the presidency of the Society, a post which has had, in the main, a distinguished list of occupants.

Coming to America in 1920 as Professor of Psychology at Harvard University, Dr. McDougall was drawn into the psychical research field in this country in a still more active way. He was elected president of the American Society for Psychical Research in 1921; and later.

when the work of that Society fell below what he regarded as minimum scientific standards, he, together with Dr. Walter Franklin Prince, Dr. Elwood Worcester, and others, founded in 1925 the Boston Society for Psychic Research. Professor McDougall was by no means inactive in the research itself. While in England, he had taken part in the investigation of Professor Gilbert Murray's apparent telepathic abilities. At Harvard, he accepted an appointment to the Scientific American Committee which investigated, along with other mediums, the notorious Margery case. (This case, like all the other applicants, received a negative decision both from Dr. McDougall and from the Committee as a whole.)

When in 1930 the Sinclair telepathy experiments were reported in Upton Sinclair's *Mental Radio,* Professor McDougall made a supplementary, though brief, investigation of Mrs. Sinclair's abilities in telepathy finding that his results, although they were inferior to those earlier reported, tended to confirm the conclusions reached.

However, it was not as an experimenter that Professor McDougall made his greatest contributions to the field of parapsychology. In view of the fact of his many interests and of his maintenance of an active research program in another line, this could not be expected. It was rather in the capacity of sponsor and activator of research at the hands of others that he most strongly advanced the field. To him, more than to any other belongs the credit of introducing parapsychology to the university lecture room and laboratory.

Professor McDougall stoutly maintained that psychical research was a proper field for mature university study. This was the thesis of an address given at Clark University in 1927. Long before that, however, Professor McDougall had himself taken definite steps to implant active psychical research in the form of strictly controlled experiments into the Harvard Psychology Laboratory itself. Finding there an unused research fund contributed for psychical research, he turned this fund to use, with the help first of Gardner Murphy and later of G. H. Estabrooks, toward furthering our knowledge of the subject of telepathy. Although Murphy's results remained inconclusive, they impressed him sufficiently to warrant his continuing active interest in the field of investigation. Estabrooks was led by his results to the conclusion that telepathy occurred under the conditions of the experiment. Professor McDougall, then, may be credited with sponsoring the first research in an American university which gave evidence of a demonstration of telepathy.

In 1930 at Duke University, Professor McDougall again inspired

active investigation of telepathy and its allied phenomenon, clairvoyance, in the Psychology Laboratory of that University. From this beginning sponsored by him, the present ESP research movement has developed through a period of eight years and appears to have found a permanent place in the academic research field. While Professor McDougall has been seldom an active participant in the research of the Parapsychology Laboratory, he has been at all times an active sponsor and critic. Without him it can hardly be supposed that the research would have been possible.

Finally, when it became necessary to found this JOURNAL in order to obtain adequate publication space for research reports in the field of extra-sensory perception, Professor McDougall lent his name as well as his active cooperation to the editorship of the new publication. The fact that this was done at the age of 65 years, a time when his own research had become something of a burden and when demands from many quarters were constantly pressing upon him, will testify to the strength of his interest and to his belief in the importance of the problems dealt with.

Those who have read his works, whether or not they have altogether shared his views, have felt the unusual breadth and force of his great intellect; while in those who have worked with him through the years there have grown up such feelings of mingled affection and respect as must be rare indeed.

EDITORIAL COMMENTS

As the JOURNAL OF PARAPSYCHOLOGY closes its second year, there is reason for thinking that the year 1938 may be a significant one in the history of the subject. It is not that 1938 turned out the highest average score in the history of the research; the Riess series of December, 1937, still holds that record. Nor is it that the research of the past year saw unprecedented experimental conditions. The Pearce-Pratt distance work, the Warner case, and Riess again, have hardly been surpassed on safeguards, if indeed they can be.

Nor is 1938 marked by the largest number of reports supporting the case for extra-sensory perception; that too came last year. Oddly enough, too, the longest series of tests in history, the Pegram work, was reported in 1937.

The introduction of the precognition experiments might appear to be this year's primary claim to importance in the history of ESP. But the necessarily slow-moving series of articles on precognition have still left us uncertain as to what the final outcome will be, and the end of this series rather than its beginning will determine its importance in the history of parapsychology.

Much importance is also being attached to the demonstration (first by Price and Pratt) of the role of the subject-experimenter relationship and its bearing upon negative series. Other contributions stand out in the year's advance and from one point of view or another are of first importance. But from the viewpoint of the progress of the research program as a whole, the event which seems above all to characterize the year 1938 is the definite emergence of this research on the stage of psychological attention. This emergence has been sufficiently disturbing to leave no doubt about its occurrence. The subjects of telepathy and clairvoyance have at last after more than fifty years of investigation come unmistakably to the attention of psychologists. This has largely occurred this year.

It is true, the research in ESP has been preponderantly in the hands of psychologists during the past twenty years. This is not only true of the reports appearing in the JOURNAL OF PARAPSYCHOLOGY but of the earlier investigations such as those of Brugmans and Heymans, at

Groningen, of Troland, Murphy, and Estabrooks at Harvard and of Coover at Stanford. All were psychologists. Many psychological leaders, too, have taken active interest in or made contributions to the field of extra-sensory perception (for example, William James, Alfred Lehmann, Pierre Janet, Sigmund Freud, and William McDougall). Nevertheless, the focus of psychological thinking was never appreciably turned to the problems of extra-sensory perception until this year.

The first evidence that psychologists were paying serious attention to the ESP work came from the spring meetings of the various sectional psychological groups; at New York, Knoxville, Madison, and on the Pacific coast, the discussion of ESP was conspicuous. Although this discussion was predominantly critical, there were some experimental reports and some constructive criticisms. The final evidence of serious consideration, however, came with the scheduling of a symposium at the general meeting of the American Psychological Association at Columbus, Ohio, on the 9th of September of this year.

The impression given there was that the majority of the American psychologists are open-minded to a presentation of the facts bearing upon the hypothesis of extra-sensory perception. That they are may presumably be inferred from the exceptionally large attendance given to the ESP symposium. If confirmation was needed, it was furnished by the generous distribution of the applause and the thoroughly fair attitude shown by the audience. No more heartening event has occurred in the history of the research and perhaps none more significant of its opportunities for progress. This year has, in that respect, made history for ESP.

Sixty-two years earlier, Professor William Barrett presented a paper to the British Association for the Advancement of Science reporting results of experiments which seemingly involved both hypnosis and extra-sensory perception. He was ridiculed and his paper excluded from the proceedings. In the 1938 meeting at Columbus there were symposia on both topics and there was no ridicule. Slow as it is, there has been some progress.

* * *

It is evident from a glance over the contents of the present number that the more purely experimental articles are being unduly crowded out by the expansion of critical reviews, questionnaires, and symposium reports. This is exactly what is occurring in the research itself. However, only a small part of the task of the pioneer investigator is actual discovery itself. The task of achieving a hearing and some measure of

acceptance often proves to be, for the early stages of a research, the much more difficult task.

It is important to many readers to know what psychologists in general think about the research in extra-sensory perception if only because of the fact that in many practical ways the explorations are themselves affected by such thinking. For example, the research personnel is itself thus affected. There are few people starting their careers in research who can afford to be entirely indifferent to the way in which their research will be regarded. Entirely apart from the question of courage, there are other very important considerations. A graduate student, or a young psychologist at no stage can easily afford to select a topic for investigation, no matter how strong his personal interest in it, which is regarded as nonsense by the majority of the men in his field. We welcome, therefore, the reports on the questionnaires of Warner and Clark and that of Crumbaugh for the light they actually throw upon the important question of what psychologists are thinking about ESP—all the more so because we have received letters from psychologists who hold no brief for ESP protesting against a third recent questionnaire which, they hold, is misleading because of the peculiar wording and leading character of its questions.

Two constructive trends of suggestions will be seen to arise in the report of the ESP Symposium given in this number. The first of these is the emphasis placed upon the advantage of independent recording in the test procedures. While a considerable amount of the work done has already had the extra safeguard of independent recording, it seems highly desirable to extend the proportion of studies which have been thus safeguarded. A second step in this direction will be represented in the several testing machines in different stages of construction in this country which will completely mechanize the selection, recording, and other features of the experiments; and doubtless such machines may prove to have other advantages as well. Progress, however, in the development of these machines to the point of satisfying every conceivable criticism in advance is necessarily slow. An endless stream of problems seems to arise because of the many demands made upon such machines. Beyond question, the development of an adequate machine merits help from all sides, both on the basis of its importance and on the strength of its difficulties.

The other constructive suggestion arising out of the Symposium discussion is concerned with the JOURNAL OF PARAPSYCHOLOGY itself. The suggestion is that space be given for supplementary statement by the authors of certain articles which would more clearly facilitate the

repetition of their experiments under identical conditions. We shall await and welcome such specific suggestions as can be made toward this end. Letters to the editors offering constructive suggestions would be in order, whether for publication or not.

* * *

So many alternative hypotheses have to be considered in reaching a tentative conclusion regarding evidence for extra-sensory perception that the reader may well be a little bewildered by the tangle of hypotheses with which an experimental report must cope. Are sensory cues adequately excluded? Has selection of the data been avoided? Has the recording and checking been done with the proper accuracy? Do the results run the gauntlet of all the various statistical criteria which must be met before significance can be conceded? These are only the main alternatives. To what extent have they been met in the experiments reported herein? Both the work by Price comparing the blind with seeing subjects and the work by Martin and Stribic reporting a series of 1,000 runs by a single subject with the unique characteristic of an increase in average as the tests proceeded apparently do well in meeting the requirements; but whether they have met all that may be asked may depend on how recently the work was done. New requirements are constantly appearing. However, even if there may in any case yet be demanded still further safeguards, the conclusion tentatively indicated may help greatly to guide the further research itself.

The question whether the blind possess superior powers of perception in other than the usual sensory channels is attacked in a systematic experimental fashion for the first time in the Price report. If, as is likely, a number of people have held beliefs on this question that are not borne out by the present research, it is hoped that someone will be spurred to an independent experimental study of the problem.

Nothing speaks more definitely for the complete voluntary control of ESP than the Martin and Stribic work with the UT condition; that is, with the subject calling up through instead of down through the pack. With such voluntary control, we ask, "Why can we not go further and subject the entire process to more volitional guidance?" It will be interesting to see what future research will yield on this question.

* * *

Many students of ESP have longed for something less mathematical and less concerned with rigid controls than the general run of experiments repeated in this JOURNAL. Others have wished to have the tests

based on more intrinsically interesting material than ESP cards. They can now be referred to M. René Warcollier's new book, *Experimental Telepathy*[1] recently published under the editorship of Dr. Gardner Murphy. The research reported in this book is based on the use in telepathy tests of "free material" (i.e. freely selected objects or drawings used by the sender) and there is no quantitative evaluation given. The explanation is entirely devoted to the discovery of insights that may lead to understanding and control of telepathy.

From Dr. Murphy's introductory paragraph we may draw an excellent glimpse of the author's purpose and the importance Dr. Murphy attaches to Warcollier's methods of approach:

"For some twenty-five years a French engineer, René Warcollier, has been at work in an effort to ascertain exactly what happens in those experiences which are commonly called telepathic. Almost alone among investigators, he has concerned himself not with the question of marshaling masses of evidence, but with the question of the mental processes which characterize those who 'send' and those who 'receive' impressions. In many phases of science it is necessary to prove up to the hilt that a certain kind of event does actually occur before one discusses how or why it occurs; in the present field, however, it is Warcollier's opinion that until we understand the how and why so as *to produce the phenomena at will,* we shall not be able to give the subject true scientific status."

* * *

Another book dealing with experimental extra-sensory perception has recently appeared on the other side of the Atlantic by an experimenter whose name has frequently appeared in the pages of this JOURNAL, in one instance as author. We refer to *Science and Psychical Phenomena*[2] by G. N. M. Tyrrell. Mr. Tyrrell, as his title indicates, includes a great deal more in his survey besides the research in extra-sensory perception, but in view of the fact that his own research has been mainly in extra-sensory perception it is upon this branch that we may look upon his book as most authoritative. His discussion of conditions favoring and hindering extra-sensory perception is most helpful, partly because we know so little about this aspect of research in this field. On page 109 Tyrrell draws a very appropriate comparison between some of the difficulties of demonstrating ESP and the difficulties of demonstrating a sense of humor.

[1] This was published by the Boston Society for Psychic Research. A somewhat abridged edition was published by Harpers under the title *Experiments in Telepathy.*

[2] Published by Methuen & Company, London, 1938. American edition, Harpers.

THE *ESP* SYMPOSIUM AT THE A. P. A.

ABSTRACT: On September 9th of this year at the meeting of the American Psychological Association at Columbus, Ohio, there was held a symposium on the subject of "Methods in ESP Research," with Dr. John L. Kennedy as chairman. Six short papers were read concentrating attention on three main issues of controversy over ESP: the methods of evaluation of results, exclusion of sensory cues, and the avoidance of errors in recording and handling of the data. The speakers were Drs. Louis D. Goodfellow, T. N. E. Greville, H. O. Gulliksen, J. B. Rhine, John L. Kennedy, and Gardner Murphy. Following these papers the meeting was opened for general discussion. The substance of the papers and the discussion is given in the following report which is based upon stenographic records of the symposium and upon copies of papers made available by the authors.

THE FIRST symposium dealing with ESP ever held at a general meeting of the American Psychological Association took place at Columbus, Ohio, at 4:00 P.M. on September 9, 1938 under the chairmanship of Dr. John L. Kennedy, who has held, since the fall of 1937, the Fellowship in Psychical Research at Stanford. The audience was unusually large for a symposium, numbering by conservative estimate over 500.

The chairman opened the meeting with the statement that the topic was "Methods in ESP Research," adding a few remarks on the importance of the question and the need of investigators and critics to reach, if possible, agreement on the procedures that should be used.

GOODFELLOW ON THE ZENITH TESTS

The first speaker introduced was Dr. Louis D. Goodfellow of Northwestern who had for a time collaborated with the Zenith Foundation in the handling of their radio-telepathy test results. Goodfellow endeavored in his paper to show that while the results of the tests were highly significant, they could be adequately accounted for without the hypothesis of telepathy. They were, he argued, explainable as the result of pattern-habits on the part of the public aided by favoring leads given in the instructions. Since this paper is published in full in the December JOURNAL OF EXPERIMENTAL PSYCHOLOGY and in any case concerns a phase of controversy the background of which has not as

yet been brought before the readers of this JOURNAL, it will not be reported here in any further detail.

GREVILLE ON ESP AND MATHEMATICS

The second paper was read by Dr. Thomas N. E. Greville, Department of Applied Mathematics and Actuarial Science, University of Michigan, on the mathematical problems of the ESP research. After introducing the questions included in the scope of his paper, Greville read as follows:

"Do Mathematical Computations Possess Evidential Value? Those critics who assert that nothing can be proved by mathematics—that the existence of ESP can be proved only by some striking demonstration of telepathy ignore the fact that such demonstrations have often been given, (as brought out, for example, in Dr. Murphy's recent article in the *American Scholar*[1]), and have usually been explained on the ground of chance or coincidence. It is difficult to see how the chance hypothesis can be met except by recourse to the mathematical theory of probabilities. The particular merit of the technique used in the recent investigations by Rhine and others is that it lends itself to quantitative treatment.

"Nevertheless, it must be recognized that the theory of probability cannot assert the impossibility that a given occurrence or set of occurrences is due to chance. It can only say that such an occurrence can be expected to happen only once in so many hundred, or thousand, or million times. Highly improbable events may and do sometimes occur, but when the same highly improbable event occurs repeatedly, a mathematician would inevitably reject chance as an explanation. Taking for illustration the well-known Pearce-Pratt series, a score as high as that obtained in Series A would be expected to occur about once in 10^{17} times. That may have just happened to be the one particular time in that large number that such a thing chose to happen; but then immediately after we have in Series B a second performance which itself would happen but once in a hundred million times. After two such "runs of luck" we should expect a normal or below chance series, but instead Series C exhibits a similar order of high scoring with a probability of once in 10,000 times. Those who have worked with ESP cards know how seldom very high scores are obtained in empirical chance series. In Greenwood's huge empirical series of 20,000 runs, the highest score ever obtained was 14. Yet in the Riess series of 74 runs we have one score of 25, three of 24, one of 23, and three of 22.

[1] MURPHY, GARDNER. Dr. Rhine and the mind's eye. *American Scholar*, 1938, 7, 189-200.

By chance alone, a single score of 25 (ignoring the rest of this dazzzling array of scores) would happen but once in six hundred trillion times. It is trite to say that the number of decks which have been run in the whole history of investigations with ESP cards could not even remotely approach such a figure. How can chance explain such results?

"Is the Theory of Probability Sufficiently Established? It is sometimes said that the principles of probability are a matter of controversy among mathematicians, so as to cast doubt on their application to scientific problems. While, just as in every other branch of science, there are differences of opinion on certain phases of the theory of probability, particularly its more philosophical aspects, I am not aware of any disagreement as to its basic methods. Certainly they have long been used extensively and without hesitation in all branches of science. Moreover, the oldest and most well established of the principles of probability are those which apply to games of chance and similar situations, such as those under discussion here.

"While the theorems of probability, like all other mathematical theorems, are of the "if . . . then" type, the only assumptions made, insofar as they apply to the present discussion, are essentially these: that in a series of thorough shuffling of a deck of ESP cards there is no reason for expecting any particular order of the cards to appear more frequently than any other, and there is no systematic correspondence between a given order and the preceding or following one. These assumptions commend themselves to the common sense of most persons, and their truth has been indicated by a large number of empirical checks.

"Has Probability Theory Been Properly Applied in ESP Investigations? It should be clearly recognized that, given the data of a particular experiment, the probability that a score as high or higher would be obtained under chance conditions is a definite, determinate quantity, for which a mathematical formula can be written. However, in most instances, the time and labor involved in making an exact calculation would be prohibitive, and it is necessary in practice to employ some approximation. The approximation most generally used is that commonly known as the probability integral. The validity of this integral as an approximation to probabilities arising in games of chance and similar situations was first demonstrated by Laplace in 1812, and has never, to my knowledge, been questioned. It is possible to fix definite (and usually very small) limits to the degree of accuracy secured in a given case.[2]

[2] USPENSKY, J. V. Introduction to Mathematical Probability. New York: McGraw-Hill Book Co., 1937. Chapter VII, especially p. 129.

"It should be recognized that the problem is definitely one of probability—not of statistics. The failure to recognize this has caused much misunderstanding and confusion of thought in the articles of a mathematical nature which have appeared on this subject in the psychological journals. A typical example is Kellogg's advocacy of the use of the standard deviation of the subject's own scores rather than that derived from the theoretical chance distribution, a wholly unjustifiable procedure. This confusion has been heightened by the fact that the actual arithmetic which must be carried out in evaluating the above-mentioned approximation is identical with certain statistical procedures; and it is unfortunate that the ESP investigators themselves have encouraged the same misunderstanding by placing so much emphasis on the so-called "critical ratio" as a criterion of significance, instead of regarding it as merely a step in the computation of the probability with which it is associated.

"I consider it unnecessary to refer to the many mathematical criticisms of a trivial nature, most of them wholly irrelevant, which were made during the early period of the controversy on this subject. It is sufficient to point out that even those which are relevant are negligible in their practical effect on the resulting probabilities. I do not know of any mathematician of repute who has seriously questioned any of the recent investigations on mathematical grounds.

"Is the Work of a Particular Subject Evidential? If the calling of ESP cards is a chance matter, the subject is only an instrument, and there is no reason to expect the work of one subject to differ materially from that of another. If the results obtained with a particular pair of dice were found to be significantly different (in the sense of a small probability of chance occurrence) from those expected on the basis of the general theory and those obtained with other dice, we should have no hesitation in concluding that these dice were loaded. In an analogous ESP situation, we should be equally justified in concluding that the particular subject was "loaded"; and if it could be demonstrated that no "loading" existed in the form of sensory cues or recording errors, the ESP hypothesis would be indicated.

"The work of a particular subject drawn from a larger collection of similar work can be separately evaluated. This problem, however, can take a considerable number of forms, and it is not possible to discuss in this brief space the various adaptations needed. Some of these sampling problems are considered empirically by Greenwood.[3] By way

[3] GREENWOOD, J. A. An empirical investigation of some sampling problems. J. PARAPSYCHOL., 1938, 2, 222-230.

of a general remark, however, the probability for the total work of an individual subject can be given a numerical estimate quite as definitely as a given block of data taken on any other basis; for example, it may be stated that one subject in 150 would be expected on the basis of chance to produce a series of results of the character obtained. This question has not, to my knowledge, been acutely raised by the ESP investigations, because the work of single subjects has not been isolated for the purpose of establishing the case. Such evaluation, however, is frequently advantageous for comparative purposes.

"Can Scores Be Inflated by Stopping Subjects at a Favorable Time? It is sometimes argued that by discontinuing work with each of a number of subjects at a time when his score is slightly above the chance level, the group average will be artificially inflated. It must be reiterated that, on the chance hypothesis, the subject is merely an instrument, and changing subjects could no more affect the score than using a different pair of dice or a new deck of cards, or walking around the table. To put the matter in another way, if a random series of calls is matched with a series of actual decks of cards and the average total score recomputed after each call, the score will naturally exhibit fluctuations above and below the chance expectation. Now, if the entire series is to be broken up into a number of subseries, the grand average cannot possibly be affected by whether the divisions are made at points where the score is above or below chance. Changing subjects could reasonably affect the score only if there is some difference (either temporary or permanent) between the two subjects which operates selectively in the direction of high scores. I cannot conceive what such a difference could be other than increased capacity for responding to sensory or extra-sensory cues.

"Miscellaneous Objections. 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 acccuracy, the probability that a given score will ever be reached in a series of infinite length. The formula can be modified so as to exclude the case where the given score is obtained in the first so many calls. This modification is desirable, since a short run is more likely to produce high scores and is not usually regarded as reliable evidence of ESP. 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.

"There seems to be a minority of psychologists and laymen who maintain that cards do not actually behave in the way that the laws of probability assert, and some claim even to have made actual experiments which bear out their contentions. To the mathematician such results are as startling as ESP itself, and apart from their connection with this subject, are worthy of careful investigation. They are certainly at variance with the results of empirical checks carried out many times over a period of years, and should be accepted with the utmost caution until confirmed independently under conditions which exclude the operation of extra-chance factors."

GULLIKSEN ON SENSORY CUES

Third on the program was Dr. Harold O. Gulliksen, Board of Examinations, University of Chicago, speaking on the subject of sensory cues in ESP tests. After a brief introduction he launched into a criticism of the cards in use in ESP tests:

"Then persons not associated with Duke University who were working on ESP discovered and published the fact that the ESP figures appear as slight depressions on the back of the card. This may be due to local shrinkage of the card upon the drying of the ink, or to the pressure of stamping the figures on the card or to both. With a little practice, and under correct lighting conditions almost anyone can attain 60 to 80% accuracy, instead of the 20% accuracy that would be expected from chance guessing of the ESP deck. What is more important is Kennedy's discovery that a subject may use such cues and be entirely unconscious of that fact. He has reported one such case in the *Journal of Psychology,* a subject who made good scores when allowed to see the back of the card and chance scores when prevented from seeing the back of the card. The subject, when questioned as to why he looked at the backs of the cards, said that otherwise "There would be nothing to go on," but was not able to make the reason more explicit. Similarly Coover showed that subjects may react to subliminal cues of which they were unconscious.

"It is also important to note that reading the cards from the back is possible not only with the commercially produced decks of ESP cards but also with the special hand-stamped decks of Zener cards like those used in the tests reported in Rhine's monograph. I have seen one such deck and found that a number of the cards could easily be read from

the back, and Kennedy also reports tests with one of the decks of Zener cards showing that they could be read from the back.

"I have a letter from Dr. Rhine in which he states that the cards I refer to as "hand-stamped Zener cards" are not Zener cards and are not those used in the earlier successful tests. Since Dr. Rhine is to follow me on this program, I will let him clear this matter up when he has the floor.

"There are possible sensory cues from two other sources, although to my knowledge no one has reported any evidence bearing directly upon these cues.

"It is possible that the figures may be identified when the subject is allowed to touch the card, either by feeling the slight depressions on the back of the card, or by the roughness of the ink marks on the front. This would require rather delicate tactile discrimination, and as far as I know no one has tried to see whether or not this cue can be successfully used under any of the conventional conditions for ESP experiments.

"Kennedy has reported some recent work on shuffling of ESP cards which suggests that kinaesthetic cues might be used also. He has shown that ordinary cutting and shuffling does not turn up the five different cards in chance proportion, but that cutting the deck with a paper knife does give a chance frequency of the different cards. This suggests, as Kennedy has said, that there may be "differential slipping and suction factors between the cards." One type of card may move more easily than another on shuffling and cutting. If these differences exist, then when the subject is allowed to handle the deck as in the various matching techniques and others, he may well detect these differences kinaesthetically. Again no one has to my knowledge done any work to test whether subjects' scores when they may handle but not see the cards are better than when they neither handle nor see the cards.

"From a psychological standpoint the possible ability of subjects to react to such obscure cues is extremely interesting; however, from the viewpoint of ESP, it is not important to investigate and prove the possible effectiveness or ineffectiveness of each cue. It is necessary only to use methods which insure that such cues are impossible.

"From what is now known it may be said that to guard against the possibility of sensory cues the experimental methods used must not allow the subject to touch or to handle the cards, and must prevent him from seeing even the backs of the cards. Several such techniques have been described in the literature, such as the DT technique (provided the back of the top card is covered), and the screened match-

ing techniques where the experimenter sorts the cards at the direction of the subject. Some experiments have been reported in which each card is kept in a separate envelope. It is difficult to see how sensory cues could operate in this case. The distance experiments also eliminate the possibility of sensory cues from the cards.

"Results obtained by the BT method may be due to the use of sensory cues as is also true of both the open and blind matching, and screened matching if the subject sorts the cards.

"Using the list of ESP methods as given by Rhine and by Stuart and Pratt one cannot classify them into those that do and do not permit sensory cues, because many of the methods are described with variants, some of which permit and some exclude sensory cues.

"It may seem to some persons that this emphasis on obscure sensory cues is paying too much attention to what is obviously a minor point in research. However it must be remembered that the bulk of the positive work on ESP shows scores that are on the average only one or two hits above a chance score in each run of twenty-five, so that the cues would need to work on only one or two cards out of the twenty-five in order to produce the results that are reported.

"In the December, 1937, issue of the JOURNAL OF PARAPSYCHOLOGY, Dr. Rhine has summarized the results of experiments that gave adequate safeguards against sensory cues, and found that the number of extra-chance hits is still significant but is lower than when the subject may use these cues. Many of these experiments still used careless methods of recording the data. The possible effect of these types of errors will be discussed by Dr. Kennedy."

RHINE ON EXCLUSION OF SENSORY CUES

As the fourth speaker on the program, J. B. Rhine also discussed the exclusion of sensory cues. He first answered Gulliksen's criticisms of the cards by explaining that neither Gulliksen nor Kennedy had observed the heavy Zener cards used in the earlier unscreened work but had instead the ESP cards provided later for use in shuffling machines, the cards to be used behind screens. It was explained also that in the first number of the JOURNAL OF PARAPSYCHOLOGY to appear after the ESP cards were put on the market, a warning was issued to use screens with the cards in all serious experimentation. He went on as follows:

"Dr. Gulliksen has generously recognized that there are three conditions; namely, screening, sealed opaque envelopes, and distance, which he is willing to accept as excluding sensory cues, regardless of those

defects of the cards to which he refers. In other words, if the subject has absolutely no sensory contact with the cards, backs, faces, or edges throughout the period of the experimental run, and if the experimenter who is handling the cards is himself not looking at the cards and is therefore unable to furnish even unconscious sensory cues, such as "involuntary whispering," then we can all agree that sensory cues under such conditions are eliminated from the experiment and we can go on to consider the improvement of conditions on other points. Dr. Gulliksen has referred to the fact that a considerable amount of work has been done under these three conditions which are acceptable to both of us.

"The exclusion of sensory cues is not, however, all that is involved in a satisfactory ESP experiment. Dr. Gulliksen rightly called attention to the question coming up later on the program; namely, recording errors. It seems appropriate for me to raise still further questions (since we must leave recording errors to Dr. Kennedy and Dr. Murphy and since we have agreed essentially upon the conditions excluding sensory cues). In a word, what other provisions must we add to those which exclude sensory cues in providing completely adequate conditions? I shall review a number of those which have seemed most important.

"Along with the exclusion of sensory cues must go, of course, adequate presentation of all the data. In other words, selection of improper character must be ruled out. We have made strenuous efforts to see that all the data obtained under the conditions described are presented. It was possible to present the first three year's work at Duke University absolutely in full. Last December, I summarized for the JOURNAL OF PARAPSYCHOLOGY all of the tests of which I had any knowledge (thus avoiding selection of results) which had been conducted under a specified set of conditions which were regarded as adequate to the exclusion of sensory cues; namely, the three conditions just mentioned, screening, using sealed opaque envelopes and of interposing distance between the cards and subject. The total of these trials numbered, with a slight correction introduced in the next succeeding number of the JOURNAL, 145,325. This gives a critical ratio of 29.5 with a probability of approximately 10 with an exponent of minus 195. I have had the totals for the three conditions put on the board. You may see from this little table that large critical ratios are evident for each of the three conditions excluding sensory cues. The point to emphasize here is that so far as we know this is a complete presentation up to the time of publication of all the data of all the subjects of all the investigations which met those conditions.

TABLE I

TOTALS OF ALL ESP CLAIRVOYANT TYPE TESTS UNDER CONDITIONS OF COMPLETE ABSENCE OF SENSORY CONTACT UP TO DECEMBER, 1937

Condition	Trials	+ Deviation	Av. hits per 25	S. D.	C. R.	P.
Cards sealed in opaque envelopes....	16,900	555	5.8	53	10.5	.....
Cards behind opaque screens.....	94,200	3,054	5.8	125.3	24.4	.....
Distance separating cards and subjects...	34,225	987	5.8	75.5	13.1	.....
Total.....	145,325	4,597	5.8	155.5	29.5	10^{-191}

"Should we bring the report up to date, we should have to add a still larger number that have been reported since which would meet the conditions described. These results would greatly increase the critical ratio and give a still smaller probability that the results are the product of chance.

"If it be suggested that perhaps we have selected the very conditions themselves to favor the high scores, I will mention that if we were to include in this group the results that were excluded by the strictness of the conditions, the only effect would be to increase enormously the critical ratio obtained.

"As a matter of fact, the highest all time record made in ESP tests has actually been excluded from this particular group. I refer to the Riess series referred to earlier by Dr. Greville, in which an average of over 18 hits per 25 was obtained through 1,850 trials or 74 runs through the pack. Dr. Riess's work was omitted from the table here on the board on the grounds that he was looking at each symbol at the time the subject was forming an impression of its identity. However, he was located at a distance of over a block from the subject, and it is clear that no sensory cues could have been used as the basis of communication between subject and experimenter. If we included all the work that has been done under this condition, admitting the possibility of a telepathic factor, but with adequate distance between experimenter and subject the critical ratio would still be increased greatly above the figure given. The critical ratio for Dr. Riess's work alone is 52.

"Thus, sensory cues, including the Hansen and Lehman hypothesis of involuntary whispering, have been ruled out in the experiments reported. In all the tests summarized in the table shown (Table I), ***no one even knew the card order!*** And in the same experiments the

selection hypothesis is clearly inapplicable. Whether these results are still subject to the recording error hypothesis, we leave to the later speakers.

"Next, we might consider, along with the questions just mentioned, whether the peculiar shape or warping of the cards might lead to results favoring high scores. This could only happen if such defects led to the appearance of a certain order in more than chance frequency and if this order were taken advantage of by the subject. If this did occur it would readily be detected by the cross-check method in which the subject's calls and the order of the cards are checked against each other, not only in the usual method, but in ways in which they were not intended to be checked; that is, the calls being checked against the wrong packs of cards. Such cross checks have often been made and have never yielded anything but chance results (see Pearce-Pratt, Martin).[4]

"Is it possible that the subject's preference for certain symbols or preference for a certain order of symbols—favorite patterns, etc.—can account for high scoring? It is obvious, too, that unless the pack of cards is itself made up of unequal numbers of symbols that the more outstanding a preference, the more likely the subject's calls will approximate five. So far as favorite patterns of symbols order are concerned again the cross-check method should reveal any effect resulting from these.

"The question of the honesty of the subjects has been raised from time to time. It is always a question in any psychological experiment in which strong motivation might enter. Here, however, it is simply an aspect of the question of the adequacy of the exclusion of sensory cues: *i.e.*, whether or not it is possible for the subject, however strongly motivated, to detect what the cards are inside the opaque envelopes, behind the screen or off at a distance. If it is possible, the critic is entitled to suspect the worst. If we agree that it is impossible to get sensory cues under these best test conditions, then the question of the honesty of the subject becomes of no importance.

"Finally, we may even consider the reliability of the investigators themselves as a legitimate question in an investigation of this character. It naturally arises in a project of research in which the conclusions are sufficiently improbable to warrant considering such remote alternatives.

[4] GREENWOOD, J. A. and STUART, C. E. Mathematical techniques used in ESP research. J. PARAPSYCHOL., 1937, **1**, 206-225. (See p. 217 ff.) MARTIN, D. R. and STRIBIC, F. P. Studies in extra-sensory perception. J. PARAPSYCHOL., 1938, **2**, 23-30. Other reports have furnished similar cross-check controls.

This hypothesis is usually excluded in science by independent confirmation of the original research and, as is not generally known, there have been a great many independent confirmations of the ESP work. There is, I think, no basis in history to suspect that so many investigators as have reported favorable results in ESP tests, should be found to be grossly unreliable.

"Another way of covering this point is to introduce whenever possible the two-experimenter procedure. This was done as far back as five years ago in the well-known Pearce-Pratt series conducted with a distance of 100 to 250 yards between the cards handled by the experimenter and the subject. Not only was the entire experiment conducted with independent records of calls and cards made and sent in sealed envelopes to the second investigator but during one series of 150 trials, two experimenters were present throughout each experimental period. The results obtained from this Pearce-Pratt series were highly significant. Since that time a number of other researches have been conducted with two experimenters responsible for the accuracy of the recording and handling of the results.[6]

"It would seem then that the principal questions raised have been met in the results under conditions already reported. No one, for example, has yet pointed out any acute need to go beyond the procedures followed in the Pearce-Pratt series just mentioned. Nevertheless, it is advisable to make as much progress as possible in reducing the human factors and eliminating any possible remaining uncertainties that may be left in our minds about the soundness of the procedures.

"To this end, we have been attempting to develop ESP test machines which will permit the conduct of the tests on an almost entirely mechanical basis."

Then follows a general description of the type of machine needed which may be omitted here since a more detailed account will shortly be published.[7]

Rhine continued: "Several machines of different design are in the process of development at present. I have brought one[8] with me which is not yet complete but which will give you some basis of judgment of its adequacy. I invite you to examine it and let me have any criticisms that arise.

"Naturally we do not know yet whether the ESP subject can demonstrate ability on these machines. That remains to be seen. How-

[6] Warner, Price and Pratt, MacFarland.

[7] Mimeographed copies of the detailed description may be obtained by interested readers on request.

[8] Designed by Professor Walter Seeley of the Duke Engineering School.

ever, the fact that Mr. Tyrrell in England obtained good results with a completely mechanized procedure of a similar character leads us to be hopeful.

"Finally, I think it important to say that the question of the methods for safeguarding these experiments is by no means the whole of the question of methods. In fact, before these methods are of any concern, there has to be a phenomenon which has some appearance of being extra-sensory perception. How to get this phenomenon under conditions of our test procedures is the greatest concern to many of us. Our plan has been to try to catch what looks like ESP and, gradually improving the safeguards as we go, bring the phenomenon, if possible, up to the point where there can be no question about the interpretation of the results. It is apparently necessary in most instances to use this procedure instead of that of plunging at once into the most developed and elaborate conditions and then searching for good ESP subjects under those already perfected conditions. We must, then, try to keep these general types of method in mind and improve them both. Toward this end, your help and criticism is very much needed and very much appreciated."

KENNEDY ON RECORDING ERRORS

Dr. Kennedy, as the fifth speaker, introduced the issue of recording errors and their possible range in ESP tests. He went to considerable length in his introduction to stress the possibilities of recording errors under the ESP test conditions urging that experimental conditions favored such errors most when the highest scoring occurred. Then followed a report of experimental work designed to show that motivated recorders make errors in the direction of their bias. "In this paper, experiments will be briefly described which indicate the validity of three further propositions concerning recording by the ESP methods.* They are:

(1) Unmotivated recorders make errors predominantly in the direction of increasing the hit score.

(2) ESP recorders who are highly motivated to produce either positive or negative deviations from the chance level made errors in the direction of their bias or aim.

(3) Motivated ESP recorders may make enough unnoticed errors to produce significant deviations from the chance level when the errors are left in the data.

* Dr. Kennedy, after inspection of the proof of this article, asked that it be clearly understood that his paper constituted only a brief outline of experimental work in progress, and that a more complete report is now being prepared.

"Methods and Controls. The following ESP methods were used this year in an attempt to study sources of error in recording: Open Matching, General or Undifferentiated ESP or Telepathy-Card test, Down Through or Pack Calling test, Distance (between two rooms) Telepathy Card test, and Distance (between two rooms) Pure Telepathy test. In all, 79,100 trials were collected; the corrected hit total was 15,-834; 15,820 hits were expected by chance; and the critical ratio (D/σ) was + .12. Of course, these figures are meaningless without a knowledge of the conditions of the experiments, but they are presented here to indicate the essentially chance nature of the results as a whole. Negative results, however, do not explain positive results, mainly because of the presence of the "gifted individual" hypothesis in ESP research. Negative results are explained by ESP proponents by stating that testing was not done on individuals who possess the power. Lehmann recognized this difficulty in 1896 and proposed that psychologists should look for explanations of positive results rather than resting content with verbal criticisms. Until positive results are adequately explained, verbal criticisms are as conspicuously unconvincing as the results themselves.

"Two controls to determine the accuracy of recording were applied to these ESP methods, usually without the recorder's knowledge. For each run through the pack, a different deck of cards was used. This deck was previously shuffled by a mechanical shuffler, and the order of symbols in the pack was separately recorded before the experiment. The person who used the deck was instructed to keep it in order. Runs in which the order of the deck was changed by cutting or dropping were thrown out. In the telepathy tests, an independent record was also made of the subject's calls without knowledge of the rightness or wrongness of the called symbol. In other words, both recorder-agent and independent recorder made simultaneous records of the percipient's calls. We were able, then, by having independent records of both calls and cards, to apply these to the agent's record and determine whether or not errors had been made.

"Errors which automatically increase or decrease the score may occur (1) at the time of recording or (2) when the run is checked for number of hits. Errors in recording may increase the score, have no effect, or decrease it. Errors in estimation of number of hits usually decrease the score since they involve failure to circle a hit which was correctly recorded. When the fixed pack method of obtaining a random series for guessing is used, some of these errors can be corrected by returning to the data. This is true for errors in counting the number of hits

and some recording errors where failures in the 5:5:5:5:5 frequencies of the different symbols to check are found. The only way to check on inversions of symbols in the card column and illusory errors in recording the called symbols is to make the independent record suggested above.

"*Results.* When the independent checks on recording were applied to the data reported above, the validity of the first proposition is indicated. A total of 400 errors in recording at the time of experiment was discovered. One hundred and forty-seven automatically increased the score, 216 had no effect, and 37 automatically decreased the score. In checking the data for hits after the symbols were recorded, 34 hits not circled nor counted were found.

"The second proposition gains credence from the examination of two sets of data. The first set of 105 runs or 2,625 trials was collected by the Telepathy Card method with a student as agent-recorder. The total critical ratio was + 1.29. This student was a particularly staunch opponent of ESP. Errors made at the time of recording were: positive, 1; no effect, 5; and negative, 3. The second set of only 250 trials, which was not included in the grand total, was obtained with the DT or Pack Calling method. The trials were unwitnessed, but the order of the cards was recorded before the experimental period. An elderly lady, who believed strongly in telepathy and clairvoyance and the ESP methods, tested herself. The record was handed in with 70 hits recorded. The order of the cards had been changed during the experiment. When the calls were checked with the previously recorded card symbols, the true total was found to be 59 hits. In this single set, at least 11 errors which increased the score were made.

"The validity of proposition III rests on results obtained with a modification of the Distance Telepathy Card test when the same lady was the agent-recorder and a friend was the percipient. The percipient sat in an adjoining room, made her calls vocally and the agent recorded the called symbols next to the card symbol upon which she had concentrated. An independent recorder sat in the room with the percipient and recorded the calls independently without knowledge of the card symbols. Two consecutive sets of 250 trials each will sufficiently illustrate the point of proposition III. In the first set, the record was handed in with 70 hits recorded. When checked with the independent record of calls, it was found that a total of 23 errors had been made in recording the percipient's calls, 18 increasing the score, 4 with no effect, and 1 decreasing the score. In the next set, the record was handed in with 66 hits scored. Comparison with the independent record showed

23 errors, 19 positive, 3 no effect, and 1 negative. With errors left in, the critical ratio (D/σ) for these 500 trials was 3.96, indicating clearly an extra-chance cause for the deviation. With errors removed, the critical ratio was reduced to + .11.

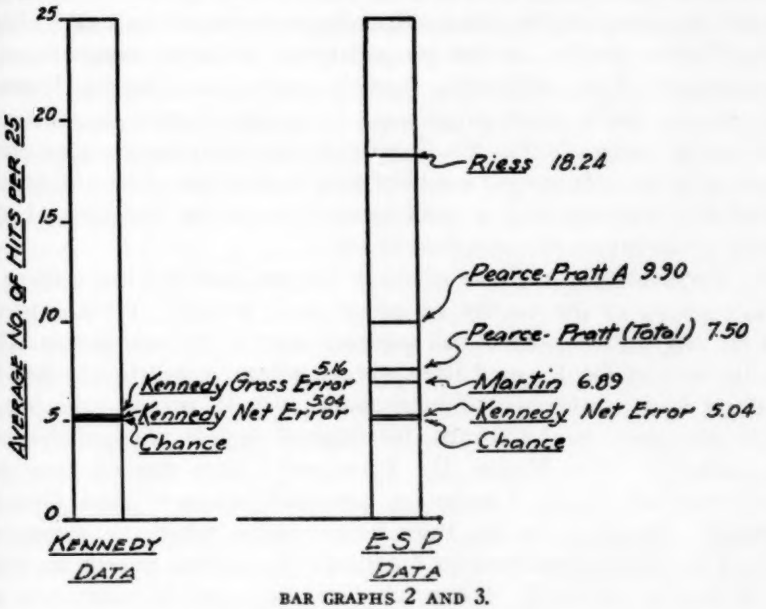
"One further experiment should be mentioned which, although it has nothing to do with recording errors, may have a bearing on the recent precognition experiments, especially the ESP shuffle. This experiment, which was performed by Mr. J. G. Moskovics, was originally intended to check the method for obtaining a random series of symbols to be guessed by shuffling and cutting a pack of commercial ESP cards for each trial. The cards were given two hand shuffles, evened up, and cut in the middle of the deck. After 5,000 trials, the chi square obtained by comparing obtained frequencies of the different symbols with expected frequencies was 87.12. To check this figure, another set of 2,000 cuts was made by a person who knew nothing about the previous results. This time, the bottom card of the cut portion was recorded in Series I; the top card of the uncut portion, which had been recorded previously, was recorded in Series II. The respective chi squares were 55.29 and 175.81, yielding probabilities that chance variation produced the results of less than .0000001. In order to partially test the hypothesis that these deviations were produced by differential slipping and suction factors between the cards, a further set of 1,000 cuts was made. The cards were shuffled in the same way but cutting was done with a paper knife. The respective chi squares for Series I and Series II in this experiment were 2.73 and 3.35 with probabilities of .7 and .5 that as bad or worse fits would be obtained by chance. These experiments seem to indicate that non-random distributions of cards in the commercial ESP cards are obtained by hand shuffling and cutting without the mediation of Extra-Sensory Perception.

MURPHY ON LIMITS OF RECORDING ERRORS

The final speaker was Dr. Gardner Murphy, Department of Psychology, Columbia University, dealing likewise with the issue of recording errors. He said that workers in Parapsychology would be grateful for Dr. Kennedy's data, with which his own unpublished material on recording errors were in substantial agreement.

These studies "indicate that when working under the specified conditions, namely without a recheck, there is likely to be one error due to recording or checking in the course of 200 trials, more or less. If these errors were all in the direction of increasing hits, we should get

the positive deviation shown in bar diagram 1. This, in a series of 1,000 trials, would mean that a critical ratio of 4.0 should be reduced to 3.7. While I have not succeeded in finding in the literature any study in which a clearly significant result would become non-significant because of such a correction, we can nevertheless all agree that such errors must be studied and eliminated. It should be added that Dr. Kennedy's data on net positive figures, *i.e.*, those in the direction of increasing hits, are as interesting as the gross total. His data, reproduced in bar diagram 2, show that in his study the net effect of errors



BAR GRAPHS 2 AND 3.

in recording is 110, *i.e.*, 27% of the 400 total. But it is of course current practice to watch for such errors both by witnessing and especially by *rechecking* with the standard deck, so that if we get more or less than the five of each symbol we can tell that an error has been made. It is indeed true that occasionally an inversion of symbols might be made, which on this basis could occur about once in the course of a thousand calls. But I did not note that Dr. Kennedy reported on the obtained frequency of such inversions nor their net effect on total hits. Calculation from his data suggests that they would come to about 20 in his series, as against 34 hits which were overlooked. It is of course necessary to work for the elimination of both types of errors.

"Actually I should imagine that the number of errors which can be made by an uncritical elderly subject would be of considerably less importance than the number made by the general run of graduate students and laboratory assistants, and I am disappointed that Dr. Kennedy did not give more data about these latter groups, especially the errors made by those students whom he reports to be biased one way or the other. The bias effect as reported by Dr. Kennedy for a graduate student is far less than enough to be statistically significant; it is a gross total of 9 errors in 2,625 trials, with a net effect of reducing the hits by 2. The biases and the errors are interesting psychological problems, but it is I think important that the effect of bias in causing errors be kept a strictly quantitative question, so that perspective on the whole situation can be maintained. I am delighted to learn in conversation that Dr. Kennedy is carrying out a much larger study of autistic factors in relation to recording. I may add that Mr. Taves and I are studying recording errors and while the data are not complete they indicate two things: first, that total errors appear to be a small fraction of one per cent; second, that most of the errors are omissions of hits.

"Fortunately in the case of much current work we can achieve an exact record of the number of errors made because of the existence of the original independent call and card records for retabulation. Thus in the case of Dr. Bernard Riess's series where, it will be recalled, the subject at a considerable distance made over 18 correct calls per 25 over the course of 1,850 calls, the original records are available here in folder R. Miss Martin, Mr. Taves and I have checked them over and found one error. I invite any interested person to check them for himself. Similarly, in the Pratt-Pearce series, where an average of over 7 was maintained for over 2,000 calls, the original records are available here in folder P. I have checked them, and in comparing my results with those of Dr. Rhine as published, find one error. I urge all those interested to do the same. As far as I know, no one has seriously suggested that the results are a statistical artifact nor that they are due to sensory cues. There are of course several other very significant series. Original records from Miss Martin are available in folder M for anyone's inspection. Bar diagram 3 aims to give an idea of the magnitude of recording errors likely according to Dr. Kennedy's data, as compared with the deviations found in various published series.

"As soon as Dr. Kennedy asked me to participate in this symposium I framed a questionnaire, the general form of which was approved both by Dr. Kennedy and by Dr. Rhine, and sent it out to every investigator who had published a study in the JOURNAL OF PARAPSYCHOLOGY. I

wanted to find out whether extensive data existed as to the likelihood of various kinds of recording errors, and especially whether the investigators agreed in general with Greenwood's principles appearing in the March, 1938 number of the JOURNAL OF PARAPSYCHOLOGY. I think you will recall that these principles commence with the suggestion that independent records be made of calls and of target cards. My respondents are almost unanimous in saying that they are glad to indorse this principle.

"Not all experimenters believe this necessary, but all believe it desirable and are glad to follow this and all other ordinary requests which seem likely to help towards mutual understanding and the general satisfaction of critics.

"It should of course be noted that the making of independent records might mean psychologically two different procedures. In procedure number one there is only one experimenter. If it is clairvoyance that he is investigating, he makes the record of the target cards after the record of the call cards. It is conceivable that he might, even when carefully witnessed, be led into an occasional error through the fact that he had previously observed the order of the calls. In procedure number two this difficulty is removed by using two experimenters. Experimenter No. 1 observes the cards or the subjects; then after the calls have been made he takes the cards to a different place beyond the range of sensory cues, and experimenter No. 2 records the target cards. The juxtaposing of the two records permits as many independent checkings as are subsequently required. Among my respondents there is general willingness to carry out this procedure for part of the current work, but also a feeling that it is hardly wise to declare on a priori grounds that no work should be performed except when two experimenters are available. There are various other ways of meeting reasonable demands, one of the simplest of which is the practice, during long distance work, of having experimenter and subject transmit independent records to unbiased third parties. I have already obtained two professional psychologists who are willing to carry out this sort of independent checking, namely R. C. Tryon, of the University of California, and R. H. Gundlach, of the University of Washington, and I hope to receive the names of others in various sections of the country. There is no time to report here the replies of all respondents to all phases of my questionnaire, but I do wish to report the prevailing opinion, in which Dr. Kennedy and I also agree, that if the independent records are made, all the other Greenwood principles are of secondary importance.

"It will be of interest I think to the present gathering that the independent record procedure was adopted at Duke last spring, and that a considerable amount of such work will soon be available in the JOURNAL OF PARAPSYCHOLOGY, in addition to the samples to which I have referred in folders R, P, and M now present for your inspection.

"Regarding the Moskovics study of hand cutting versus penknife cutting, Dr. Kennedy's report leaves me uncertain regarding several points. Frankly I wish that Dr. Kennedy had performed this experiment, or had repeated it himself, or had told us more about the method. Whether suction or slipping factors are responsible for the results I do not know. Of course the usual problem in cutting cards is not whether the card above and the card below the cut are in some way linked in a non-chance way, but whether the obtained order of the cut deck agrees beyond chance with another deck. I should be much interested in such an investigation. Most investigators, I am sure, will be very glad to cut with penknife rather than by hand if this will satisfy other experimenters; yet I think that the best control method in the study of extra-chance relations between decks will probably continue to be the customary cross-check from call cards to target cards.

"One of the most interesting of the hypotheses with which Dr. Kennedy commenced his paper was this: 'The experimental conditions under which ESP appear at its height are those which foster the making and "channeling" of errors in recording.' This is indeed a crucial question for an empirical testing. It seems appropriate therefore to look through the published studies and select the one which actually shows ESP at its height. This proves to be the Riess series, which was done at long distance with the independent record method. The next step would seem to be to look for the second highest average ESP score, and this proves to be the Pearce-Pratt series, also done by the method of independent recording. I think perhaps Dr. Kennedy was responding to his understanding of certain phrases regarding the necessity for a casual attitude on the part of the subject, and was very naturally misled into a generalization which is not well supported by inductive procedure."

GENERAL DISCUSSION

Dr. Kennedy opened the meeting for discussion.

Mr. Hyman Rogosin (New York University):

I want to say first that I had imagined, from the way in which the limitations of this discussion had been presented, that perhaps philosophy of science as a means of arriving at correct conclusions in psychology and in science in general, had been excluded from this round table. Yet I think that

this phase of the work has a very definite place in regard to experimental techniques. And so I am glad that Mr. Greville in his paper took up a number of points in which I am very much interested, and mentioned specifically recent views in mathematics and psychology. Now, mathematical support is not all in favor of the methods used and elaborated by the workers in ESP. For example, I have before me a letter from Prof. R. A. Fisher of England, stating: "Perhaps I may say, with respect to the use of statements of very long odds, that I have before now, criticised their cogency on the grounds, not only that the procedure of calculation is often questionable, but that they are much less relevant to the establishment of the facts of nature than would be a demonstration of the reliable reproducibility of the phenomena." This, I think, attacks one of the fundamental points made on mathematics by the Duke University people. The issue I want to present tonight is the one constantly pointed out in mathematics: be careful of the assumptions you make; be sure that you know the psychological assumptions in regard to mathematics in using certain material; and especially should you be careful in dealing with probability theory. It is a controversial subject, but that is not the main point. The main point is that in dealing with this particular material you should be extremely careful in drawing your interpretations and conclusions. Mathematics cannot *in itself* determine claims; mathematics *alone* cannot prove anything. It is absolutely essential in mathematical studies to present the qualifications that are necessary; in ESP work, it is extremely important to mention and heed the qualifications that most writers on probability point out. On that ground I want to say that there have been careless ways of approaching the problem and of reasoning in this whole field of ESP.

Dr. Herbert C. Sanborn (Vanderbilt University):

Which are they?

Mr. Rogosin:

I haven't time for them here. I have published two papers in which I point out the errors and criticisms I have found. There will be at least three more published shortly, and I will be glad to give their place of publication to anyone interested.

Dr. Knight Dunlap (University of California at Los Angeles):

Mr. Chairman, I should like to ask Mr. Greville a mathematical question bearing on the selection of data. Mr. Greville pointed out that it is perfectly proper to stop the work of a particular subject as soon as it drops below the chance level. To make it a little clearer in my mind, I'd like to offer this analogy. Suppose I were to throw a pair of dice and try for a seven (a bad number!), and they turn seven the first or second throw. So instead of keeping on throwing for sevens with the first pair of dice I take a second pair. But they run pretty badly so I don't count them. The third pair of dice don't give any better scores so I throw those out—they don't count and I take a fourth pair. Suppose we go on that way, only counting the scores of the dice giving the best results. Would you say the percentage

cannot be artificially raised by using only the favorable material, just as if you choose only to count the scores of the subjects which make the best records?

Dr. Greville:

I think it quite obvious that if we counted only the scores of the dice giving favorable scores an artificially inflated result would be obtained. I am not going to state just what methods were used in conducting the various published experiments. However, as I understand it, all the scores obtained, favorable or not, were recorded. This question might best be answered by those directly concerned with the investigations.

Miss Dorothy Martin (University of Colorado):

Mr. Chairman, may I at this point present evidence on the question just raised. The question whether it is possible to obtain artificially high scores by selecting only the favorable data has been answered by Dr. Murphy who referred to the chart on the right which represents the work of a single subject over a period of nearly two years. The total number of runs is 2,500. The total number of trials is the rather stupendous figure of 62,500, which I think you will grant is a rather long series. Every score, high or low, was counted and included in the data of the published report. The subject began scoring rather low, 5.6, and showed definite improvement until the June vacation of the year 1937. Note that when I began again in September there was another drop which rapidly increased until the 7.8 average was reached. This series covers more than a thousand runs. Again the Christmas vacation intervenes and again we note the score is lower but gradually increases to around 7.2 or 7.4. I offer this in proof of the fact that this subject has not been stopped when he was scoring high, that no records have been thrown away, and further that all the records I have ever taken with this boy are reported here. He is an undergraduate, a willing student—subject, I should say. The last 20,000 trials were made under the most carefully controlled conditions, all using screens, many with a third observer.

Dr. Steuart Henderson Britt (George Washington University):

Dr. Kennedy, I should like to speak of a point which I think will settle this controversy, and I am expressing a point of view which a number of people have given to me. That is, one should not rush into print before having all results and conclusions fully stated. I observed that Dr. Rhine spoke of "hypothetical extra-sensory perception." But you will recall that the book *Extra-Sensory Perception* was not called "hypothetical extra-sensory perception," that the conclusion given in that book was that extra-sensory perception had been proved.

Now I wish to call attention to the fact that the conditions have been changed repeatedly as reported in that monograph. In the first series the subject had contact with the cards, helped to score the cards, and helped to check his own records. Then someone pointed out that sensory cues might be obtained from having the subject handle the cards. Dr. Rhine became aware of that error and said, "All right, we will scrap that data." Where-

upon the subjects were not allowed to touch the cards and opaque screens were used, and this material was presented saying that it proved ESP. But later on it was decided these were unsatisfactory and Dr. Rhine says that material should be scrapped; so, that doesn't prove ESP. So now we introduce into the picture wooden screens and distance. More data are presented showing that this is ESP. But again more data are introduced showing that that was not ESP.

I simply want to call attention to the fact that material is introduced again and again which bolsters up the data, based on large numbers of trials with conditions which have been scrapped such as sensory cues, shuffling of the deck by the subjects, helping score their own records, and the use of opaque screens. So I simply raise the question of whether we are at all justified in talking about experimental procedures when actually all Dr. Rhine can say is, "Here is a hypothetical something or other that we are in search of." We cannot judge any of the results obtained in working over the data.

I repeat—the fundamental proposition in any field of science is that we should have all our results reported, know all the implications of our data, and be sure that all experimental problems are fully worked out before making any claims.

Dr. Lucien Warner (New York City):

I think Dr. Britt has actually given us in brief outline the history of this research. I just want to say that probably the work would never have gone any further if the earliest results were not published. The book was responsible for creating interest in a questionable subject. As a consequence the next work was done better and attracted still greater attention. It is the rule in all the sciences that research makes headway not through withholding all work until perfection is attained but through the give and take of criticism and further improvement in later experiment. That is the only way we are going to have progress. As soon as there is such a situation as the Pratt-Pearce series I think it is justified to publish that and then go on to the next instead of waiting to see that every critical condition has been met.

Dr. J. G. Pratt (Duke University):

Mr. Chairman, I am somewhat puzzled by the attitude of the speaker preceding the last in criticizing Dr. Rhine for varying the conditions when presumably we have met here to discuss means for improving the methods in ESP. However, I think that this audience will have observed that in the papers that were read there was a marked division of subject matter between the so-called opponents and proponents of the ESP research. The critical papers of the opponents have stopped short at certain series which offer the best conditions. The other papers have taken brief account of the poorer conditions and have gone on to take fuller account of conditions which have not been criticized; I might say, conditions which have so far defied criticism.

I should like to say on my own authority that the examples shown on the chart before us can be called Exhibit A on the question "Does ESP

occur?" That does not mean that there are not experiments which may be subject to other possible explanations, but it seems to me that in a symposium where methods are being considered we should consider first the best methods and point out just what is wrong with them. We may find that they are good enough to satisfy everyone concerned. At any rate, they show how far the methods have already progressed, and further improvement must take off from the best experimental conditions if it is to represent any advance in methods at all. The critics seem to be doubtful about this issue. We would appreciate such criticism (and thus I identify myself with the proponents of ESP) as can be made on these better experimental conditions.

To change the subject briefly, I am surprised that the ghost of "selection" has continually hovered over these meetings. This is particularly surprising when we note that the committee arranging this session did not consider it necessary to have anyone to read a paper criticizing the mathematics. But I think we might get some comfort on this question of selection if we consider from our Exhibit A the Pearce-Pratt series total. That has an extra-chance value of 10^{27} as evaluated by one method. As evaluated by other methods the value approximates that measure of improbability. I have had my pencil busy here for a moment, and I find that if eighteen billion people each made a million chance comparisons of ESP cards for an infinite time, they could expect to find by their combined efforts an average of one such series per billion years. Now I dare to suggest that this one small fragment of Exhibit A could not have been found by chance even if selection had been practiced deliberately, and I wonder why the ghost of selection still lives.

Dr. Kennedy:

I have been asked by Dr. Murphy to call upon Dr. Riess of Hunter College to tell us about the conditions of his experiment.

Dr. Bernard F. Riess (Hunter College):

There is nothing much to say. This subject worked for me a while last year, made 74 runs and then stopped because of her condition of health. She had a sort of general breakdown. She has since, unfortunately, disappeared into the Middle West and at present is not available for further work. So far as errors in recording are concerned, I hear Dr. Murphy found one error, and I owe him ten dollars for finding that one. But with that error eliminated, the only other error that may have crept in is a possibility of deception, and the only person who could have done the deceiving is myself since the subject at no time knew how well she was doing nor had any idea of the cards which were being turned by myself in my room, a quarter of a mile away from where she was working. Outside of that there can be no criticism of the method used. I had the deck of cards on my desk, shuffled them and at the stated time turned them over one by one making a record of each card. I had my desk piled with examination papers and half the time I don't think I was conscious of any serious concentration on my part. I kept the records locked up in my desk and sometimes it was a week before I totaled up the scores and found the number of high scores she was making. Incidentally the whole thing started when a group of students

at Hunter College wanted to do some of the experiments. I undertook the experiment as a way of demonstrating to my classes that ESP did not occur. I did not succeed in that. I had 67 other students who worked for me, none of whom had an average score above 5. So I was left with this one case. After I got the data collected and found that the subject averaged around 18 I did not know what to do but publish the results. I do not know whether Dr. Britt believes in throwing away good data just because he doesn't precisely understand the full implications of those data, but I felt that it should be reported.

Dr. A. Brett (University of Tennessee):

May I suggest that subject C. W. J. (Miss Martin's subject at the University of Colorado) be sent around to some other universities to be tested?

Dr. S. B. Sells (U. S. Works Progress Administration):

Mr. Chairman, many of the comments made today are concerned with possible errors or lack of control in published reports. During the last summer I attempted to review the literature critically. I found it very difficult to reconstruct the experiments on ESP reported in the JOURNAL OF PARAPSYCHOLOGY and other journals from the conditions reported. I find certain crucial data have been omitted. I think all the data and all the conditions should be stated: what method of testing was used, how the cards were handled, how the records were kept, and so forth. I wonder whether Dr. Rhine would consider it possible to set aside a few pages of the JOURNAL OF PARAPSYCHOLOGY and invite contributors who have published previously to submit the necessary data to make the published account of their data complete enough to reconstruct the experiment from beginning to end?

Dr. Rhine requested a repetition of the question.

Dr. Sells:

I asked whether there could be set aside a certain amount of space in the JOURNAL OF PARAPSYCHOLOGY for the purpose of having authors, who have already published, furnish further information about the procedures and conditions of their experiments which were omitted from the original reports. I might give one or two examples of the difficulties I have encountered. In some papers I could not tell which cards had been used although they said they had used "regular ESP cards." Other reports omitted details concerning the exact set-up of the experiments or of the mathematics used. I feel that if these additional facts could be presented, it might be possible to make an evaluation of the papers that have already been published.

Dr. Rhine:

Well, I will begin—and I hope I have the authority—by appointing Dr. Sells as chairman of a committee, to be extended later, to furnish us with specifications not only upon past publications, but to become a current critic of the JOURNAL; and, if he has time and patience and is willing, to keep shooting in to us any criticism he thinks, if taken up properly, would im-

prove the publication. I think that we need help on the JOURNAL just as much as we do on the experimentation, and I would appreciate knowing of anyone else who is willing to join this committee. If he will let me know, we will send him copies of the JOURNAL—if the committee does not get to be too large. All we have been able to do is to use our own inexperienced judgments—if we leave out Professor McDougall's—on these problems and we shall appreciate any help we can get.

Dr. Kennedy:

The hour is so late that we must close this discussion. I suggest, however, that if there is interest in continuing the discussion tomorrow morning a time and place will be arranged and posted on the bulletin board in the administration building.

(It was agreed later that it was not convenient to meet further, and an appropriate notice was posted.)

A COMPARISON OF BLIND AND SEEING SUBJECTS IN *ESP* TESTS

MARGARET M. PRICE

ABSTRACT: A group of blind subjects were compared with a corresponding group of seeing subjects for performance in ESP tests. The two groups were mainly boys resident in institutions; the one a school for the blind, the other an orphanage. Similarity in choice of subjects, test conditions, and number of tests was approximated as far as possible. Sixty-six blind and 40 seeing subjects were tested and a total of 52,975 trials were given with an average of 5.61 hits per 25. The results as a whole were highly significant giving a critical ratio of 14.06. The blind subjects gave slightly but fairly consistently higher average scores than the non-blind. About one-fourth of the runs were made with cards enclosed in sealed opaque envelopes; and in both groups, these gave significantly higher results than did the open cards. About one-third of the total runs were made by the matching methods. These gave a slightly higher average score than with the calling procedure.

Miss Price was, at the time of the completion of this work, a graduate student in the Department of Psychology of Duke University and a research assistant in the Parapsychology Laboratory. In conjunction with Miss Margaret Pegram she published, in June, 1937, a report of an investigation of ESP with blind subjects.

A PREVIOUS investigation was reported[1] in which blind subjects were found to give highly significant results in tests for extra-sensory perception. An outstanding feature of this earlier investigation was the fact that a higher percentage of subjects gave significant deviations than had been reported for seeing subjects. Whether this constituted evidence that ESP capacity was more developed among the blind could not be determined on the basis of that study alone. No investigation had been made of a comparable institutionalized group of non-blind subjects. To provide a basis for comparison an investigation was projected among institutionalized boys of age and intelligence range similar to that of the blind subjects.

SUBJECTS

Blind group. The blind subjects tested in the earlier research were boys enrolled in the State School for the Blind in Raleigh, North Caro-

[1] PRICE, M. M. and PEGRAM, M. H. Extra-sensory perception among the blind. J. PARAPSYCHOL., 1937, **1**, 143-155.

lina (the criterion for entrance is inability to read pica type at a distance of one foot from the eyes). Ages ranged from 8 to 23 years. The superintendent of the school selected the 61 subjects as a representative group. There were, in addition, five subjects used who were employed in the Durham Workshop for the Blind. These ranged in age from 19 to 35 years.

These 66 subjects varied educationally from first grade pupils to high school graduates. In degree of blindness, too, there was variation; there were a large number of partially blind subjects with wide differences in the amount of sight. Nine of the total number, however, had been entirely blind from birth; five more had become so during childhood.

Seeing group. The seeing subjects were boys enrolled in the Oxford Orphanage, Oxford, North Carolina.[2] They ranged in age from 8 to 16 years. Their school classification ranged from the fourth to the eleventh grade. The 40 subjects used were selected by the superintendent simply as representative children and without any reference to imagined ESP ability.

The two groups were the most nearly similar in age, educational level, and institutional life that it was possible under the circumstances to obtain. They seem sufficiently similar in constitution to allow a comparison of capacity for ESP.

METHODS

ESP tests were given with two types of card packs. One was the regular ESP pack, consisting either of the ESP cards shown in the frontispiece of this JOURNAL, March, 1937, or of the present commercial ESP cards. The tests with the blind were made with the former and those with the seeing with the latter. The second set consisted of ESP cards sealed in envelopes. The sealed packs were made by pasting a blank card over the face of an ESP card, sealing this double card in a brown coin envelope, and taping the outside of the envelope with broad strips of gummed paper. A code known only to the experimenter was written on the outside of the envelope. In order to insure the subject's having no opportunity to learn this code, the sealed cards were never checked within the sight of the subject but behind an opaque screen. (See photograph illustrating screen.)

In order to acquaint the blind subjects with the symbols, five sample cards with raised symbols, made by pasting figures cut from

[2] The courtesy and cooperation of the institutions which aided in this work are gratefully acknowledged.

PLATE 1 (above): Illustrating use of note-book as screen in BT tests given blind subjects.

PLATE 2 (above): Illustrating screen used in BT tests with Oxford subjects. Cards, even when sealed, are handled and responses recorded out of sight of the subject.

PLATE 3: Blind children competing in a BM test with sealed cards. The order of the key cards (in center) is unknown to the subjects, who match the unknown sealed cards to the key cards.

cardboard on the cards, were presented to each subject before the tests were made.

Test Procedures with the Blind Subjects. For GESP (General Extra-Sensory Perception), the cards were shuffled and the pack was placed face down on the table before the subject but screened from his view. The experimenter removed the top card and looked at it while the subject attempted to call it. The card was then laid aside and the next card treated the same way. This was continued for the 25 calls. The screen concealed the cards but not the experimenter. (See photograph.)[3] This was found to be the most satisfactory method for introducing subjects to the experimental procedures because these subjects seemed to think it more "reasonable" to try to name the cards when the experimenter was looking at them. This method was largely confined to this introductory use. As in all procedures used, no indication of success or failure was given until the end of the run of 25 calls.

A shift to BT (Before Touching) was usually made with each subject near the end of the first experimental period. BT is similar to GESP except that the experimenter does not look at the cards until after the 25 calls have been made. Logically, in GESP both clairvoyance and telepathy are possible; in BT telepathy is presumably ruled out. In the BT tests, the cards were screened as in GESP.

Two matching techniques were used: (1) OM (Open Matching), (2) BM (Blind Matching). All the matching techniques have one feature in common: the subject makes his guesses by distributing the twenty-five target cards in five different piles. For OM and BM there are five key cards consisting of one each of the five symbols. These five cards are arranged in a row in front of the subject and he indicates his guess for each target card (of the pack in his hand) by placing it opposite the appropriate one of the five key cards.

In OM the key cards are placed face up and the subject is told the order of the cards. Cards with raised symbols were used as key cards for some of the subjects, particularly the younger subjects who were totally blind, so that they might touch the raised figures while the experimenter placed the cards opposite the key card indicated. This procedure was discontinued when the subject became thoroughly familiar with the order of the key cards.

BM differed from OM in that the key cards were placed face down, so that the subject attempted to match symbols whose order he had not been told and could not know himself through the senses against key

[3] The screen used was a 9" x 12" notebook open and standing on end, with the cards close behind it.

cards that were likewise unknown. With this procedure, the cards with the raised symbols were never used.

A few runs were made with a procedure called Match Piling or MP. For MP no key cards were used. The subject distributed the 25 cards in five piles and then, after all the cards had been placed he "named" the piles; that is, he touched each pile and said "This is the circle pile, that is the star pile," etc. These results are included with BM which the method most resembles (and the results of which they approximate in score average).

Sealed card packs were used in great part with all the methods except for GESP, which was done entirely with open card packs.

Test Procedures with the Seeing Subjects. The test procedures used with the seeing subjects were essentially those used with the blind. The differences were: (1) A larger screen (16" x 24"; see photograph) was used in BT tests with the sealed cards than was used with blind subjects. (2) When unenclosed cards were used the subject was allowed to witness the checking and recording. (3) Very few GESP runs were made, and these only to introduce the subject to the tests.

With the seeing subjects most of the Open Matching tests were conducted with the sealed pack. A few runs were made with the regular pack with the 16" x 24" screen before the subject so that he could not see the cards he was handling. All Blind Matching and Match Piling tests were conducted with the sealed cards.

THE EXPERIMENTAL SITUATION

The atmosphere during the whole experiment was kept as informal as possible. The subjects viewed the experiment somewhat as a game in which each wished to do as well as he could. Candy rewards were occasionally given, but they were used infrequently so that they did not come to be goals in themselves. That is, the subjects did not as a rule expect or work for the reward. Competition was used when particular subjects seemed to respond well to it.

In the tests of the blind subjects, the boys were allowed to stand around while they were awaiting their turns and to talk among themselves as long as they did not interfere with the subject or seem to be having a disturbing effect on him. With the conditions of freedom that prevailed, it was necessary to have certain strictly enforced rules. (1) No person was allowed to stand beside or behind the experimenter. (2) For all GESP and BT work, a screen was used. The cards were held behind and close to it by the experimenter so that they

were invisible to the subjects and spectators while the calls were being made. (3) The subject was kept at least four feet from the experimenter, usually seated at the opposite side of the table.

Additional precautions were taken even with the blind subjects. For unscreened work (i.e., for conditions other than BT and GESP), sealed packs were used in great part. The subjects were never told what symbol card was in any envelope. The code was never revealed to any subject; the code numbers were so small as to be indistinguishable by any blind subject. Three different packs of sealed cards with three different codes were used, and the sealed packs were checked behind the screen.

Similar normal precautions were taken with the seeing subjects, except that there was seldom a third person in the room. Whenever a second subject was present he was seated with the other across the table from the experimenter, and the two were allowed to work either alternately or competitively. The screening used in all this work was wide enough (24″) so that neither subject could see around it.

Recording and Checking. All of the results submitted for analysis in this paper are based upon tests in which there was a record made both of the call series and of the cards. In addition a certain amount of the testing was done with the matching methods in which only the score was recorded. Although these results will be mentioned below, they will not be used as the basis for conclusions.

The recording was done on the standard ESP record sheets recording call series and card series in parallel columns. Following the recording of the cards, each hit or coincidence was checked by encircling it and the number of hits recorded.

In order to determine whether the recording of the card series was in any way influenced by the proximity of the call series, a check was made to discover whether there was any departure from the even number of symbols (five each) expected in the card column. Since the standard pack contains five of each of the symbols, any departure from this might reasonably be supposed to be either an error in recording or an error in the composition of the pack. Only those runs are reported in which there is a perfect check obtained by this method.

It is recognized that if two errors are made in recording a run, they might conceivably cancel each other so far as this method of checking is concerned. However, since the cards were turned up and recorded one at a time and the danger of inversion of symbols thus removed, any such accidental inversion would be expected to occur

only once in 25 such runs in which errors were made. Danger from this source then may well be dismissed from consideration.[4]

RESULTS

The analyses and comparisons submitted in this report are based upon a total of 2,119 runs or 52,975 single trials combining the results of tests given to both blind and seeing subjects, but consisting only of results which have been checked and rechecked by the methods described above. As will be seen by reference to Table I, these 2,119 runs gave a positive deviation of 1,296 and an average of 5.61 hits per 25. The critical ratio is 14.06; and the probability of chance occurrence about 10^{-44}.

The purpose in limiting this comparative study to scores capable of being rechecked was to exclude directly the hypothesis of recording error. However, since at the beginning of the experimental work the value of such a recheck was not foreseen many runs were made which could not be completely rechecked. The following accounting of these runs is made in order to complete the record of work done and especially to show that in excluding these runs from the comparison there was no unwitting exclusion of low scores.

There were 930 runs made which could not be rechecked. They averaged 5.65 hits per run (as compared with 5.61 for the rechecked work). Of these, 396 runs, with an average of 5.66, could not be rechecked because of the absence of records of either card or call series or both. The remaining 534 runs consisted of those made with ESP decks in which the cards were not equally distributed as to suit. The accuracy of the card series records could therefore not be determined by a recheck. All these unchecked runs are eliminated from consideration in the conclusions of this report. (On the hypothesis of recording error, however, it is interesting to note that if such errors occurred in this work, they must have occurred in those eliminated runs, yet the average is only .04 of a hit per run higher than the rechecked scores. Any motivational trend in the hypothetical errors must have been negligible.)

Comparison of blind and seeing subjects. Keeping entirely within the results of the rechecked series described above, it is found that the 2,119 runs consist of 1,387 runs made with the blind subjects and of 732 made with the non-blind. The totals are not exactly comparable

[4] The rechecking of the results recorded in this paper was made by Mr. E. P. Gibson, record librarian of the Parapsychology Laboratory and additional sample rechecks were made by Mr. Burke M. Smith.

TABLE I

TOTALS FOR ALL SUBJECTS AND CONDITIONS

	Trials	Runs	Deviation	Average	S. D.	C. R.
Rechecked.....	52,975	2,119	+1,296	5.61	92.06	14.06
Unrechecked....	23,250	930	+ 603	5.65	60.99	9.83
Total.....	76,225	3,049	+1,899	5.62	110.43	17.22

because the work with the seeing subjects was interrupted before its completion.[5] It will hardly be disputed, however, that the control series of 732 runs or 18,300 individual trials represents an adequate basis for comparative judgment. The distribution of these runs over the various types of procedures followed in the test is roughly comparable in both series, that with the blind and with the seeing subjects. A comparison of the averages, therefore, obtained from these two series, should throw some light upon the ESP capacities of the two groups of subjects compared; at least of their capacity to perform under these conditions.

It will be seen from Table II that while the 40 non-blind subjects produced an average of 5.54 for the 18,300 trials, the blind subjects gave an average of 5.64 in the 34,675 trials. The difference between these averages is not in itself significant by statistical measurement. It is of interest, however, to see that the superiority of the blind holds consistently through the various subdivisions of these series that are indicated in Tables II, III, and IV below.

The exceptions in Table V occur only with small numbers of runs; therefore, even though the difference favoring the blind is small, it may be safe to attach some importance to it because of this consistency of distribution. (It should be added that this difference in favor of the blind does not persist in the unchecked series.)

Comparison of open and enclosed cards. In the earlier report on the work with blind subjects mentioned above, a higher average was reported for the results of tests in which cards enclosed in opaque sealed envelopes were used than resulted from tests with open cards. A similar difference in favor of the use of enclosed cards was reported by J. B. Rhine.[6] These differences reported were small, consisting of .39 in the first instance and .30 in the second.

[5] The circumstances of interruption (the ending of school) were not optional and accordingly the "optional stopping" modification recommended by Greenwood (J. PARAPSYCHOL., Sept. 1938) is not applicable.

[6] RHINE, J. B. ESP tests with enclosed cards. J. PARAPSYCHOL., 1938, **2**, 199-216.

TABLE II

COMPARISON OF BLIND AND SEEING SUBJECTS

Subjects	No. of Subjects	Trials	Runs	Deviation	Average	S. D.	C. R.
Blind.....	66	34,675	1,387	+ 895	5.64	74.48	11.93
Seeing.....	40	18,300	732	+ 401	5.54	54.10	7.41
Total....	106	52,975	2,119	+1,296	5.61	92.06	14.06

A comparison of the results of the enclosed and the open card tests is offered in Table III, based entirely, as stated above, upon the rechecked series. This last feature accounts for the difference between the figures earlier reported for the blind subjects and those found in Table II. The earlier figures, however, gave a difference of .39 between the sealed and the open cards and the present table gives a difference of .36. Compared with this the averages for the seeing subjects differ by .41 in favor of the enclosed. It is suggestive to find such relatively similar figures resulting incidentally from analyses of these results based upon widely different groups, obtained by different experimenters, and indicating some advantage in performance with the cards enclosed in sealed opaque envelopes. Furthermore, the difference between the averages for the total of the open and enclosed card tests submitted in Table III is statistically significant.

TABLE III

COMPARISON OPEN AND ENCLOSED CARDS

Subjects	Open					Enclosed				
	Runs	Dev.	Ave.	S.D.	C.R.	Runs	Dev.	Ave.	S.D.	C.R.
Blind.....	1,069	+603	5.56	65.36	9.23	318	+292	5.92	35.66	7.95
Seeing.....	505	+213	5.42	44.94	4.74	227	+188	5.83	30.13	6.24
Total.....	1,574	+816	5.52	79.34	10.28	545	+480	5.88	46.68	10.27

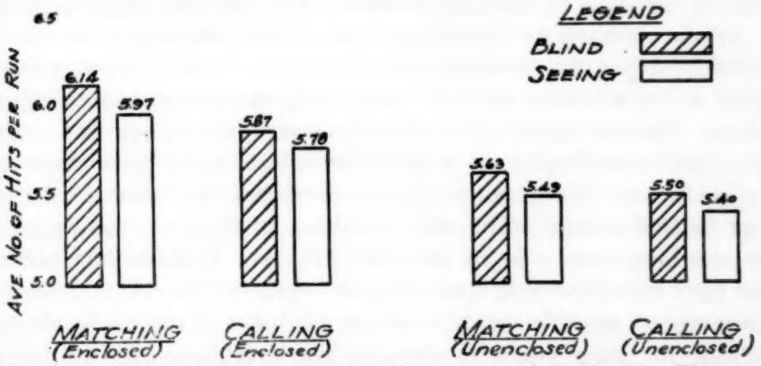
Comparison of matching and calling procedures. Keeping in mind both of the earlier comparisons, first, that of the blind and the seeing, and second, that of the open and enclosed cards, it is of interest to compare under these headings the results of the matching and calling procedures used in the tests. The matching tests are characterized by a manual mode of response and the calling by a vocal one. Under matching are included the open matching and blind matching; and under

TABLE IV

COMPARISON OF MATCHING AND CALLING PROCEDURES

Subjects	Matching					Calling				
	Runs	Dev.	Ave.	S.D.		Runs	Dev.	Ave.	S.D.	C.R.
Blind	483	+304	5.63	43.95	6.92	586	+299	5.50	48.41	6.18
Seeing	117	+ 58	5.49	21.63	2.68	388	+155	5.40	39.39	3.94
Total	600	+362	5.60	48.98	7.38	974	+454	5.47	62.41	7.27
	ENCLOSED									
Blind	57	+ 65	6.14	15.09	4.31	261	+227	5.87	32.31	7.03
Seeing	61	+ 59	5.97	15.62	3.77	166	+129	5.78	25.76	5.02
Total	118	+124	6.05	21.72	5.72	427	+356	5.83	41.32	8.62
Grand Total	718	+486	5.68	53.58	9.07	1,401	+810	5.58	74.86	10.82

calling, the BT and GESP. Both with the open and the enclosed card series with both blind and seeing subjects, the matching procedure yielded slightly higher averages than did the calling procedures. These differences, however, are in no instance significant. Further details may be found in Table IV and in Graph I. It will be seen that in these subdivisions of the data the relations indicated above persist throughout: namely, the blind performance was superior to that of the seeing subjects; the enclosed cards series gave better results than the open; and the matching procedure excelled over the calling. The only one of the three comparisons that gave a statistically significant difference, however, was that resulting in favor of enclosed cards over open.



GRAPH I. Representing in diagram the results of Table V. The words "matching" and "calling" describe the test procedures; "enclosed" and "unenclosed" refer to the cards used in the tests.

DISCUSSION

Unless in this report we have to do with acceptable evidence for extra-sensory perception the comparisons made from the results are meaningless. It is necessary then to review the issues that naturally arise when we consider whether we have to do with evidence for the hypothesis of extra-sensory perception. These issues will be considered under the headings of evaluative, experimental, and personal.

The critical ratio of 14.06 given by the research results taken in their entirety puts the issue of the chance hypothesis entirely aside. Unmistakably, something meaningful is happening in the response of the subject to these tests. The fact, too, that all the results of all of the subjects in all of the tests are included in Table I can leave no question of selection of data. The right to exclude from consideration in the comparisons of this report the results which could not be rechecked will not be questioned statistically.

Subjects were taken as they became available without any basis for a priori judgment of their probable success. Subjects were tested for their appointed test period and were not dismissed from further testing because of low scoring; though again this would have been permissible statistically.

On the experimental issue, the most plausible alternative hypothesis to that of extra-sensory perception would be that of the Hansen and Lehman hypothesis of involuntary whispering. This would apply however, only when the experimenter knew what the cards were and presumably might unconsciously convey the impressions to the subjects by auditory cues. Only one of the methods followed permitted knowledge of the cards by the experimenter. This was the GESP procedure. It will be observed by examining Table V that this procedure was followed in only a very small group of the tests. These GESP tests were given as a preliminary measure introducing the subject to the tests as a whole. For the other procedures, it is not easy to see how sensory cues could have functioned to assist the subject in his performance. If the cards were not sealed in opaque envelopes they were in the main kept behind screens. The only exception to this rule was in part of the matching work done by the blind subjects. It is obvious, however, that their visual handicap was sufficient to prevent them from taking any advantage of possible sensory cues on the backs of the cards which are difficult for good vision to detect. The fact, too, that the subjects scored better in the tests with enclosed cards than with the open cards is argument that they were not relying upon sensory cues in these tests.

The matching tests with the seeing subjects were conducted with screens in such a position as to prevent the subjects from seeing the backs of the cards that they were handling. In the calling tests (mostly BT), both the blind and the seeing subjects worked with the cards in the hands of the experimenter and behind the screen. This situation was maintained even with the totally blind subjects and with the enclosed packs. The average of 5.87 given by blind subjects with enclosed cards held behind a screen in a series of 244 runs should leave no question as to the adequacy of the exclusion of sensory cues in at least that section of the experiment. The experimenter in handling these enclosed cards kept them with the code marks on the envelopes inverted so that she could not herself become aware of the symbol on the envelope.

In view of the precautions taken in handling the records, first, in recording both calls and card series (followed by independent checks on card frequency), and second, in rechecking hits scored and the arithmetic involved, there is perhaps little more than can be done without introducing a second experimenter. As it happens, this too was done in both the work with the blind and the work with the seeing subjects at the Oxford Orphanage. The earlier report of the work with the blind included the work of Miss Margaret H. Pegram, graduate assistant in the Parapsychology Laboratory at Duke University. Miss Pegram was not only a witness in some of the writer's tests but also found in administering similar tests to some of the subjects used that she too obtained a significant positive deviation as a result.

The following year there was instituted by Dr. J. G. Pratt and the writer a joint experiment conducted at the Oxford Orphanage requiring the presence simultaneously of both experimenters during the tests. This experiment was reported (June 1938) in two sections, the first section representing a modification of the writer's usual procedure in handling the subjects. This resulted in insignificant deviations. The second section represents work in which the writer was allowed to follow her customary policy and under these conditions the scores for the joint experiment were approximately those of the Oxford work by the writer when working alone; namely 5.52 as compared to 5.54. Both experimenters were equally responsible for the recording, checking, and handling of the data.

The hypothesis of extra-sensory perception seems the only one adequate to account for the results of the study. It is in order then to consider and discuss the comparisons reported above.

In view of the results in Table II comparing the score averages of the blind and non-blind subjects, we can hardly be greatly impressed by

the supposition that the blind possess markedly superior ability of the character of extra-sensory perception. At any rate, the order of difference, something less than 20% of the deviation of the non-blind subjects, is far from a striking one.

On the other hand, that there is a difference as consistently distributed through the subdivisions of the data as this proved to be is a fact of considerable interest and is one that, without taking too much for granted regarding its establishment, is worthy of some speculation. First, it should be pointed out that the difference, if not a chance one, may well be attributed to the writer herself. The work with the blind subjects took place a year earlier and it is not unreasonable to suppose that the experimenter herself may have changed in the meantime. The joint research with Dr. Pratt mentioned above showed certainly that there is a marked difference between experimenters and, indeed, showed that the experimenters can well obtain different results under different conditions as in the two sections of that work. So small a decline from the first work with the blind to the average obtained with the seeing subjects could well come within the scope of this possibility.

A second consideration that should be included in this discussion is the fact that the test procedures seemed to offer somewhat more entertainment to the blind subjects than to the non-blind. The curriculum for the seeing subjects offered a wider range of activities and by contrast the ESP tests were relatively less interesting than they presumably were to the blind. Interest is very probably of considerable importance.

A third suggestion is that blind subjects are less distracted by happenings in the surroundings. The lowering of the number of sensory distractions may favor extra-sensory concentration. This interpretation has been suggested to explain the slightly favored scoring of distance tests and of enclosed card tests reported by J. B. Rhine.[7] The hypothesis is that the complete removal of the cards from vision may result in the avoiding of conflict of attention between sensory and extra-sensory perception.

It seems clearly possible that any one of these three suggestions might apply to the small difference in results obtained and they may not represent all the possible hypotheses. Further research will thus be required to determine whether there be a reliable difference obtainable between the blind and the non-blind; and, if such a difference exists, whether these or other hypotheses are applicable.

[7] The effect of distance in ESP tests. J. PARAPSYCHOL., 1937, **1**, 172-184; and ESP tests with enclosed cards, J. PARAPSYCHOL., 1938, **2**, 199-216.

The apparent superiority of the enclosed card procedure is, as is indicated above, apparently in line with the superiority of blindness itself in the ESP tests. This work, together with that of J. B. Rhine mentioned above, is sufficient to warrant further pursuit of this question in its own right; namely, whether removing the object from all sensory contact effectively favors perception by extra-sensory means. The work with the blind, with distance, and with enclosed cards points strongly in this direction. On the other hand, motivated factors are not controllable in such tests and may be complicated. For example, when the enclosed cards were presented to the blind it was with a special challenge, the experimenter herself showing greater interest in the task. This was done to offset anticipated decline of interest in the subjects and overcome a feeling of futility in the face of the greater barrier the enclosing offered.

The slight superiority of matching over calling procedures recalls the fact that Woodruff and George[8] and Gibson[9] both reported higher results for open matching than with calling procedures, but had still lower results with blind matching. In both researches the combined matching methods would have averaged below the combined calling methods. There is, then, a difference in the present work in that combined matching methods are superior to combined calling methods and as will be seen from the data of Table V there is not the marked difference between blind and open matching which was reported by the other experimenters.

TABLE V

COMPARISON OF RESULTS OF SPECIFIC TEST PROCEDURES

Subjects	Open								Sealed							
	BT		GESP		OM		BM etc.		BT		GESP		OM		BM etc.	
	Runs	Ave.	Runs	Ave.	Runs	Ave.	Runs	Ave.	Runs	Ave.	Runs	Ave.	Runs	Ave.	Runs	Ave.
Blind.....	513	5.47	73	5.79	237	5.59	246	5.66	244	5.87	17	5.82	28	5.86	29	6.41
Seeing....	358	5.33	30	6.23	47	5.30	70	5.63	166	5.77	0	....	36	5.61	25	6.48
Total..	871	5.41	103	5.92	284	5.55	316	5.65	410	5.83	17	5.82	64	5.72	54	6.44

The difference in favor of matching may be attributable to the fact that the majority of the subjects in the present research were children

[8] WOODRUFF, J. L. and GEORGE, R. W. Experiments in extra-sensory perception. J. PARAPSYCHOL., 1937, **1**, 18-30.

[9] GIBSON, E. P. A study of comparative performance in several ESP procedures. J. PARAPSYCHOL., 1937, **1**, 264-275.

or adolescents. The more game-like character of the matching procedures and the greater amount of muscular action involved might safely be regarded as adding a stronger appeal to those procedures than to the calling procedures.

SUMMARY

The results of the tests for extra-sensory perception given to 66 blind subjects are compared with those given to 40 seeing subjects. Both groups were selected for similarity of age and institutional status. The total results are highly significant, giving a critical ratio of 14.06.

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

The averages obtained by both groups were higher when enclosed cards were used in the tests than when ordinary open cards were used. Both for the blind and the non-blind and both for the enclosed and unenclosed cards the score averages with the matching procedures were slightly higher than those of the calling procedures.

STUDIES IN EXTRA-SENSORY PERCEPTION: II. AN ANALYSIS OF A SECOND SERIES OF 25,000 TRIALS

DOROTHY R. MARTIN and FRANCES P. STRIBIC

ABSTRACT: A long series of trials (25,000 in number) were performed under uniform conditions, by a single subject. The up through or UT technique was used throughout. All tests were screened. An average of 7.39 with a critical ratio of 37.01 was obtained.

Two control series were employed: (1) the reverse series, in which the guesses of the subject were checked against the reverse or down through order of the pack; (2) the matching series in which each pack sequence was checked against the pack sequence consecutive to it. These series yielded averages of 5.13 and 4.91 respectively, with critical ratios of 1.95 and —1.38.

Detailed analysis of the data revealed: (1) marked consistency in daily scoring, once the establishment of the new technique had been effected; (2) greater frequency for consecutive hits in direct than in control series; (3) high positive correlation between scoring success and card position; (4) marked symbol preference without significant relation to scoring success.

A previous report by Miss Martin and Miss Stribic appeared in the March, 1938 number of this Journal.

Introduction

This is the second of a series of articles to be presented in consequence of studies made at the University of Colorado concerning problems in the field of perception without known sensory means. In the first article of this series (3) were presented results of a long series of trials (25,000 in number) performed under uniform conditions by a single subject. The DT or *down through* technique was used throughout. In this article will be presented a comparable series of trials (again 25,000 in number) performed by the same subject, using in this instance, however, the UT or *up through* technique. As in the previous report, these data, together with two control series of equal length, will be subjected to rather complete analysis.

General Conditions

The materials used throughout these tests again consisted of standard, hand-stamped (not commercial) ESP cards. These cards were arranged in packs of twenty-five, there being in each pack five cards of

each design. Chance alone would thus yield an average expectancy of five correct guesses out of each twenty-five.

The subject was Mr. C. J., who, in the previous DT series (3), averaged 6.89 hits per run. As in that series, the subject, seated at a table, was separated from the experimenter by a 15" x 18" opaque wooden screen which had no apertures. The UT or *up through* method was used exclusively. The experimenter shuffled and cut the pack, and placed it face down before her, taking care to avoid seeing the bottom card. The subject then recorded his guesses for the *up through* order of the pack. After the completion of each run, and in the presence of both subject and experimenter, the pack sequence was recorded and this recorded pack order was checked against the guessed order. The data[1] thus treated will be termed the *direct series.*

CONTROL SERIES

Two control series, obtained from variation in treatment of the original records, were introduced:

1. The *reverse series* in which the subject's guesses for each run were compared with the reverse or *down through* order of the pack, and the 'hits' noted.[2]

2. The *matching series* in which each recorded pack sequence was compared with the pack sequence consecutive to it and the coincidences noted. In order to achieve 1,000 comparisons, the last sequence was compared with the first.

RESULTS

TABLE I

SUMMARY OF RESULTS FOR DIRECT, REVERSE, AND MATCHING SERIES

Subject	Runs	Trials	Hits	Av./25	Dev.	σ*	C. R.
Mr. C. J. - *direct*....	1,000	25,000	7,387	7.39	+2,387	64.50	37.01
Mr. C. J. - *reverse*...	1,000	25,000	5,126	5.13	+ 126	64.50	1.95
Matching Series.....	1,000	25,000	4,911	4.91	— 89	64.50	— 1.38

* $\sigma = 1.02\sqrt{npq}$, the maximum standard deviation upon a matching hypothesis.

In Table I are presented the data—both *direct* and *reverse*—for 1,000 runs (25,000 trials), together with the *matching series.* It will be seen that the *direct series* yields an average of 7.39 hits per 25, with the highly significant critical ratio[3] of 37.01. The *reverse series* yields

[1] Here, as in all previous reports, *all* data obtained are included.
[2] These and all other records were doubly checked.
[3] $\text{C. R.} = \dfrac{\text{deviation from chance expectancy}}{\text{standard deviation for chance}} = \dfrac{d}{\sigma} = \dfrac{\text{h-np}}{1.02\sqrt{npq}}$

an average of 5.13[4] hits per 25, and a critical ratio of 1.95. The *matching series* yields an average of 4.91 and the clearly non-significant critical ratio of —1.38.

Further assurance of the significance of the *direct* scores, as contrasted with the *reverse* and *matching* scores, may be obtained by reference to the critical ratio of the difference[5] between these scores. Between the *direct* and the *reverse series,* this C. R. $_{\text{diff.}}$ yields the highly significant figure of 24.09; between the *direct* and the *matching series,* the similarly significant figure of 26.58; whereas the value (2.43)[6] derived from the *reverse* and the *matching series* is of doubtful significance. Attention may be called to the fact that these critical ratios, in contrast to those previously obtained, are based upon the empirical rather than upon the much attacked theoretical[7] values for the standard deviation.

TABLE II

DISTRIBUTION TABLE

Subject	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Mr. C. J. - *direct*....	0	3	14	22	50	100	159	166	176	154	78	48	21	8	0	0	1
Mr. C. J. - *reverse*...	3	27	63	122	163	209	170	125	67	35	13	2	1	0	0	0	0
Matching Series.....	0	30	75	144	184	199	173	91	61	31	7	4	1	0	0	0	0

Frequency data for the three series are recorded in Table II and are given graphic representation in Fig. 1. It is worthy of note that in the *direct series* the mode occurs at 8, whereas in both the *reverse* and in the *matching series* it occurs at 5.

ANALYSIS OF RECORDS

1. *Daily Fluctuations in Scoring.* In order to study daily fluctuations in scoring, summaries of the daily records have been compiled and are presented in Table III. Graphic representation of the daily averages for the *direct,* for the *reverse,* and for the *matching series* is given in Fig. 2. It will be noted that the *reverse* and *matching* daily averages fluctuate both above and below the chance expectancy of 5 hits per 25,

[4] Should correction be made for the excess in *reverse* hit frequencies arising from card 13 (where a *direct* hit necessarily counts also as a *reverse* hit), the average in this series would fall between the values 5.02 and 5.06, with a critical ratio in the range .23 to .99.

[5] $\text{C. R.}_{\text{diff.}} = \dfrac{M_1 - M_2}{\sigma_{\text{diff.}}} = \dfrac{M_1 - M_2}{\sqrt{\sigma_{M_1}^2 + \sigma_{M_2}^2}}$

[6] Correction for card 13 reduces this value to the range 1.17 to 1.72.

[7] Use of the theoretical values would have here yielded: 24.70, 27.68, and 2.44, respectively.

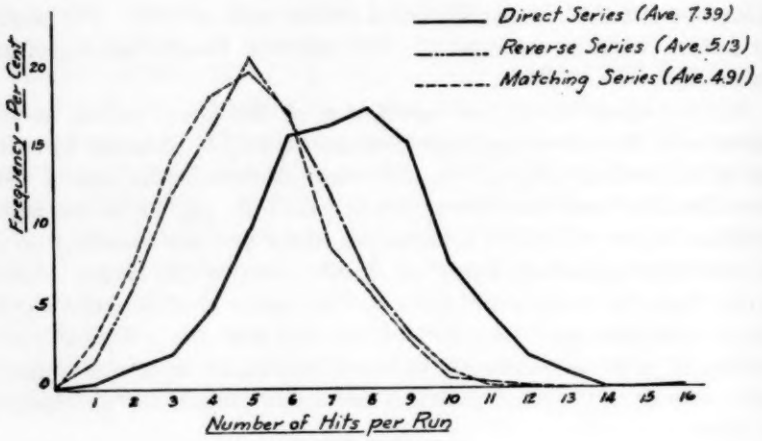


FIGURE 1

whereas the *direct* daily average remains consistently above this value. Indeed, it is noteworthy that despite change from the DT to UT technique, the subject, after an initial "warming up period," soon attained daily averages well above 7, and that these averages were, with one exception, maintained throughout the remainder of the series.

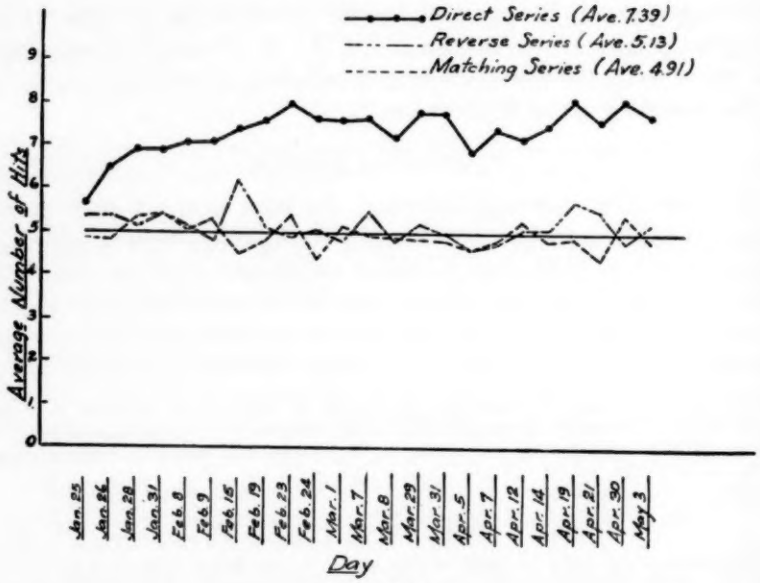


FIGURE 2

TABLE III

DAILY FLUCTUATIONS IN SCORING

Date	Runs	Direct Series			Reverse Series			Matching Series		
		Score	Av./25	C.R.	Score	Av./25	C.R.	Score	Av./25	C.R.
Jan. 25.....	50	281	5.62	2.1	268	5.36	1.2	242	4.84	— .6
Jan. 26.....	30	194	6.47	3.9	161	5.37	1.0	145	4.83	— .4
Jan. 28.....	20	138	6.90	4.2	103	5.15	.3	107	5.35	.8
Jan. 31.....	20	138	6.90	4.2	109	5.45	1.0	109	5.45	.9
Feb. 8.....	50	352	7.04	7.1	255	5.10	.3	254	5.08	.3
Feb. 9.....	20	143	7.15	4.7	96	4.80	— .4	107	5.35	.8
Feb. 15.....	20	148	7.40	5.3	125	6.25	2.7	91	4.55	— .9
Feb. 19.....	40	304	7.60	8.1	206	5.15	.5	189	4.73	— .9
Feb. 23.....	40	321	8.02	9.5	197	4.92	— .2	217	5.42	1.3
Feb. 24.....	40	305	7.62	8.2	202	5.05	.2	176	4.40	—1.9
Mar. 1.....	40	303	7.58	8.0	188	4.70	— .9	207	5.18	.5
Mar. 7.....	40	306	7.65	8.3	219	5.48	1.5	198	4.95	— .2
Mar. 8.....	40	287	7.18	6.8	193	4.83	— .5	198	4.95	— .2
Mar. 29.....	50	389	7.78	9.6	263	5.26	.9	244	4.88	— .4
Mar. 31.....	80	623	7.79	12.2	399	4.99	— .1	385	4.81	— .8
Apr. 5.....	60	412	6.87	7.1	278	4.63	—1.4	277	4.62	—1.5
Apr. 7.....	70	521	7.44	10.0	341	4.87	— .5	344	4.77	— .9
Apr. 12.....	50	363	7.26	7.8	257	5.14	.5	265	5.30	1.0
Apr. 14.....	50	374	7.48	8.6	257	5.14	.5	242	4.84	— .6
Apr. 19.....	40	325	8.13	9.8	227	5.68	2.1	197	4.93	— .3
Apr. 21.....	50	377	7.54	8.8	275	5.50	1.7	222	4.44	—1.9
Apr. 30.....	40	323	8.08	9.6	193	4.82	— .5	219	5.48	1.5
May 3.....	60	460	7.66	10.1	314	5.23	.9	286	4.77	— .9

2. *Consecutive Hits.* As has been previously noted (3), the subject when scoring well frequently scores several hits in succession. A tabulation of the frequency of occurrence of these consecutive hits, together with those in the control series, is presented in Table IV. Here again is evident a striking contrast between the *direct* and the control series.

TABLE IV

CONSECUTIVE HITS FOR DIRECT, REVERSE AND MATCHING SERIES

Series	2	3	4	5	6	7
Direct.....	1,034	345	82	45	8	5
Reverse.....	585	122	37	5	3	0
Matching.....	602	108	26	4	0	0

3. *Relation of Scores to Position of Card in the Series.* In the previous study of this series, a marked relationship ($\rho = .84 \pm .04$) was shown to exist between success in scoring and position of the card in the pack. With the change of technique from the *down through* to the *up through* calling of the pack, it became imperative to determine

whether this relationship between hit frequency and card position would be maintained or varied. Tabulation was therefore made of the frequency of scoring for the top card, the second card, the third, etc. This tabulation for the *direct,* for the *reverse,* and for the *matching series* is contained in Table V. The rank with respect to hit frequency for each of the 25 card positions is here indicated.

TABLE V

RELATION OF SCORES TO CARD POSITION

Card	Direct Series		Reverse Series		Matching Series	
	Hit Frequency	Frequency Rank	Hit Frequency	Frequency Rank	Hit Frequency	Frequency Rank
1	356	4	209	9	188	19
2	375	1	205	12	199	12
3	373	2	200	15	185	23
4	365	3	219	4	188	19
5	300	11	221	2	210	3
6	332	6	212	6	203	8.5
7	333	5	189	22	196	13
8	306	10	219	4	201	11
9	314	8	179	24	204	5.5
10	296	12	211	7	189	17
11	325	7	208	11	204	5.5
12	307	9	194	19.5	187	21
13	283	15	283	1	190	16
14	286	14	175	25	204	5.5
15	268	18	209	9	193	14
16	281	16	182	23	219	2
17	290	13	201	13.5	186	22
18	272	17	194	19.5	202	10
19	241	24	219	4	172	25
20	249	21	201	13.5	203	8.5
21	244	22	193	21	188	19
22	258	20	199	16.5	181	24
23	264	19	209	9	192	15
24	243	23	199	16.5	223	1
25	226	25	196	18	204	5.5

It will be noted: 1. that for all card positions the *direct series* frequency is above chance expectancy (200); 2. that in the *direct series,* the twelve highest ranking frequencies appear in the first twelve card positions, whereas the seven lowest ranking frequencies appear in the last seven card positions, thus suggesting, as before, a positive correlation between frequency and card position. Application of the Spearman Rank Difference Method yields $\rho = .94 \pm .02$. In the *reverse series,* the hit frequencies, with one exception, exhibit an apparently random distribution with respect to card position ($\rho = .30 \pm .13$). The unique advantage here exhibited by the middle position of card "13" is to be

expected. In the *matching series,* random distribution is definitely exhibited ($\rho = .07 \pm .14$).

4. *Symbol Preference.* In the endeavor to determine whether a symbol preference may again be exhibited by the subject, and to investigate the bearing of such preference on his scoring success, the tabulation to be found in Table VI was made. In it will be found the frequency of guesses for each symbol, the *direct* and the *reverse* hits for each symbol, and the ratio of hits to guesses.

TABLE VI
SYMBOL PREFERENCE

Symbols	Number Guessed	Hits		Ratio of Hits to Guesses (%)	
		Direct	Reverse	Direct	Reverse
Circle.....	5,547	1,637	1,159	29.51	20.89
Wave.....	5,326	1,528	1,049	28.69	19.70
Star.....	5,043	1,461	1,035	28.97	20.52
Cross.....	4,813	1,457	1,004	30.27	20.86
Rectangle.....	4,271	1,304	879	30.53	20.58

This tabulation reveals: 1. that a marked discrepancy exists between the ratio values in the *reverse* and in the *direct series,* the *reverse series* again yielding chance expectancy (20%) and the *direct series* again exceeding this expectancy; 2. that there exists but slight fluctuation in the ratio value within each series. Reference to the previous record (3) of the subject reveals for the present series not only a less pronounced symbol preference, but also a partial shift in preference, only the two least preferred symbols—the cross and the rectangle—retaining their former preference ranking. The fact that with these least preferred symbols are again associated the highest ratios of hits to guesses, suggests the need for further analysis of the symbol data. At the moment, however, symbol preference would appear to be of doubtful significance in scoring success.

Discussion

The authors regard as worthy of note:

1. The marked contrast between the results of the *direct,* the *reverse,* and the *matching series.* It will be recalled that the three series are derived as a result of subjecting identical experimental data to differing treatments. In the *direct series,* the subject's intention that his calls be matched against the *up through* order of the pack is

respected. In the *reverse series,* this intention is violated, the subject's calls being matched against the reverse or *down through* order of the pack. Clearly, if chance alone is to be credited with high scores, indifference to direction of operation must also be postulated. Actually, however, the *direct series* yields a highly significant average score (7.39), whereas the *reverse series* yields a statistically non-significant average (5.13). It becomes obvious that the only factor which could operate to effect this difference is that of the subject's intention to guess *up through* rather than *down through* the pack. In the third or *matching series,* on the other hand, each recorded pack sequence is checked against the pack sequence consecutive to it, thereby subjecting the pack data to a mechanical or "robot" calling. The results of this series not only provide a second control on the *direct* or subject series, but also, by virtue of their coincidence with or divergence from chance predictions, both theoretical (2, 4) and empirical (1, 3), serve as a useful measure of the adequacy of shuffling. The results here reported (average 4.91) establish that adequacy for this experiment. Clearly, the subject's "intention" has not functioned through a shuffling factor.

2. The maintenance, in the *direct series,* of the striking relationship between hit frequency and card position. The persistence of this relationship, despite reversal in technique, suggests that the dominant factor here involved may rest upon a physical rather than upon a psychological basis. It therefore becomes desirable to vary the experiment so as to provide for each card in turn the apparently advantageous top position. This experimental variation we hope to accomplish in the near future.

CONCLUSION

In a second long series of trials (25,000 in number), performed under the *up through* technique, an average (7.39) significantly above chance has been maintained by a single subject. Sensory cues of any known type have been excluded by the experimental controls. Comparison of the *direct* or subject series with both the *reverse* and the *matching series* establishes for this series a unique statistical position which would seem to be independent of all experimental factors except the subject's intention and ability to call the cards in the *up through* order of the pack. The results here reported, when supplemented with those of the previous *down through* study, would seem to offer strong support not only to the postulate that extra-sensory perception may exist, but also to the further postulate that it may be subject to conscious direction.

It has here been impossible to present full and detailed comparison between the data obtained in the present (*up through*) and in the previous (*down through*) experiment. The authors consider these data of sufficient interest to warrant more complete and intensive analysis. This analysis is now in progress. It is hoped that it may soon be reported.

BIBLIOGRAPHY

1. CARPENTER, C. R. & PHALEN, H. R. An experiment in card guessing. J. PARAPSYCHOL., 1937, **1**, 31-43.
2. GREVILLE, T. N. E. Exact probabilities for the matching hypothesis. J. PARAPSYCHOL., 1938, **2**, 55-59.
3. MARTIN, D. R. & STRIBIC, F. P. Studies in extra-sensory perception: I. An analysis of 25,000 trials. J. PARAPSYCHOL., 1938, **2**, 23-30.
4. STERNE, T. E. The solution of a problem in probability. *Science*, 1937, **86**, 500-501.

A SURVEY OF PSYCHOLOGICAL OPINION ON *ESP*

LUCIEN WARNER and C. C CLARK

ABSTRACT: A survey was made of the judgment of psychologists concerning the research in extra-sensory perception covering the 603 full members of the American Psychological Association. Sixty per cent replied indicating by a large majority (eighty-nine per cent) that a scientific investigation of ESP is legitimate. Seventy-six per cent included ESP research within the province of academic psychology. Thirty-nine per cent approved, in general, the experimental approach as exemplified by the best work published in this Journal.

Drs. Warner and Clark have both been contributors to earlier numbers of this Journal. Dr. Clark is Associate Professor of General Science in New York University where Dr. Warner, now engaged in private research, was one time instructor in Psychology.

A QUESTIONNAIRE survey of the opinion of psychologists regarding extra-sensory perception has been made during the past summer. The 603 full members of the American Psychological Association were interrogated by mail. Sixty per cent replied, using stamped, addressed envelopes which were provided. All but nine enclosed filled-in questionnaires, and these nine wrote at some length explaining why they had not done so. Many of the returned blanks were accompanied by long and thoughtful letters. The margins and backs of fully half the returned blanks were found to be well covered with spontaneously expressed opinions and criticisms. Many of these possess a value beyond anything which can be reflected in this brief, analytical report of the returns. We wish to express our appreciation of the interest shown in this study, and of the patience and courtesy of those who went to such trouble to make their replies as accurate and complete as possible.

An analysis of the returns is given in the accompanying table. It has been simple enough to tabulate opinion so far as it was reflected in straightforward checkings of yeses, noes, and the like. But reduction of the extensive commentary to tabular form has been difficult. In imposing classification upon this material we have endeavored not to do violence to the opinions expressed. The questionnaires have been preserved, of course, and will be made available for further study to

anyone with a legitimate interest in them. About half of the returns bore signatures, though none were requested. A very few writers requested anonymity.

In the table, the material which appeared on the printed questionnaire is italicized. The unitalicized wording represents the spontaneous contributions which were written in. Of course, this contributed material has been highly concentrated. For example, a psychologist may have written a whole page outlining an experiment planned to throw further light on ESP, and the net result in the table would be simply one more check against the suggestion that we "study the subjective and objective conditions facilitating and inhibiting ESP." Another may have given a page to his explanation of Rhine's "apparent" evidence for ESP. In the table this might be represented only by one more check opposite the demand that there be eliminated "the error of improper selection."

The questionnaire was divided into four main sections. The first requested classification of opinion of ESP under one of five suggested headings representing degrees of acceptance or rejection. Fourteen per cent believed ESP to be an impossibility. Judging by the accompanying comments, many classified it as impossible because of its definition* rather than because they believed that perception was impossible under conditions prohibiting sensory function as such function is at present understood. Thirty-six per cent preferred to check it as a "remote possibility." Several underlined the word, "remote," and several indicated that they checked this only because it would be unscientific to call anything impossible. The few who bracketed "remote possibility" and "likely possibility," or checked both, have been included under the former heading in the table. Only seven per cent believed ESP to be a likely possibility, and only five psychologists (fewer than two per cent of those replying) felt it to be an established fact. In the balance of the replies, ESP was listed as "merely an unknown." This easy way out was taken, then, by forty per cent.

The table is so arranged that the replies to the remaining questions are broken up according to the response given to this first question. For example, twenty-one of the fifty-one psychologists who believe ESP to be impossible also believe that the investigation of ESP is not a legitimate scientific undertaking. Only thirteen of the three hundred and one remaining psychologists are of that opinion.

* A brief discussion bearing on this problem is to appear shortly in the *Journal of Psychology* under the title: "Is 'Extra-sensory perception' extra sensory?"

TABLE OF RESULTS

I. *In your opinion, is "extra-sensory perception"*						
1. *An established fact?*	142	51	128	26	5	
2. *A likely possibility?*						
3. *A remote possibility?*						
4. *An impossibility?*						
5. *Merely an unknown?*						
II. *Is this opinion based on*						*Tot.*
1. *Reports in newspapers and magazines?*	45	8	39	6	3	101
2. *J. B. Rhine's book?*	65	16	66	48	5	164
3. *Experimentation as reported chiefly in the Journal of Parapsychology?*	62	14	54	11	4	145
4. *Hearsay?*	10	1	14	3	0	28
5. *Purely a priori grounds?*	41	30	45	4	0	120
My scientific background	4	6	3	0	0	13
Professional meetings	7	0	6	1	0	14
Discussion with colleagues	2	1	7	1	0	11
Reports in sources other than Journal of Parapsychology	17	5	22	5	2	51
Discussion by Rhine	3	1	0	1	0	5
Personal experience	0	0	3	2	2	7
Personal experimentation	6	4	13	5	2	30
III. *Do you consider the investigation of ESP a legitimate scientific undertaking?* Yes	134	29	120	26	5	314
No	8	21	5	0	0	34
Blank	0	1	3	0	0	4
If so, does it fall within the province of academic psychology? Yes	110	23	108	22	5	268
Partly	5	0	1	0	0	6
No	6	13	5	2	0	26
Blank and doubtful	21	15	14	2	0	52
Where else?						
Nowhere else	3	4	6	1	0	14
Nowhere at all	0	2	0	0	0	2
In an entirely new field	1	1	2	1	0	5
Physiology, neurology, etc.	3	0	7	0	1	11
Abnormal, applied psychology	0	0	1	1	0	2
Sociology	0	0	1	0	0	1
Physics	2	2	7	0	0	11
Mathematics, statistics	1	1	6	1	0	9
Philosophy	0	0	1	1	0	2
Theology, religion, mysticism	1	2	3	1	0	7
In any laboratory	1	0	2	1	0	4
In any *private* laboratory	2	3	1	0	0	6
By committee of scientists	2	0	3	1	0	6
By Societies of Psychical Research	1	1	2	0	0	4
Supported by a foundation	1	0	0	0	0	1
With cooperation of magicians	0	0	4	0	0	4
With cooperation of detectives	1	0	1	0	0	2

TABLE OF RESULTS (Continued)

IV. *In your opinion, is the present experimental movement (as exemplified by the better controlled work reported in the Journal of Parapsychology) in the proper direction?* ..... Yes	42	9	41	18	4	114
Probably	4	0	2	1	0	7
Partially	5	2	8	1	1	17
No	20	13	13	1	0	47
Probably not	1	0	1	0	0	2
Blank and doubtful	58	19	43	2	0	122
Have not read it	12	8	20	3	0	43
. . . . .	. .	. .	. .	. .	. .	. .
If not, what criticisms have you, or what methods to suggest?						
Enforce better controls, in general	14	14	11	2	1	42
Eliminate the error of improper selection	7	3	5	0	0	15
" statistical and probability error	13	1	7	1	0	22
" bias on the part of investigators	9	4	12	2	0	27
" incompetent, untrained investigators	2	3	5	0	0	10
" subject-experimenter relationship	0	1	1	0	0	2
" sensory cues	3	2	2	0	0	7
" recording error	3	0	1	0	0	4
" illusion and suggestion	2	1	0	0	0	3
" fraud	0	1	4	0	0	5
Investigate mediums	0	0	1	0	0	1
" orientals, primitive peoples	0	0	1	0	0	1
" spontaneous hunches	0	0	0	1	0	1
" the mechanism	0	0	1	1	0	2
" effect of exchanging subjects and experimenters	2	0	1	0	0	3
Abandon ESP cards; introduce other methods and material	6	2	8	1	2	19
Study subjective and objective conditions facilitating and inhibiting ESP	4	0	5	1	1	11
. . . . .	. .	. .	. .	. .	. .	. .
Unsolicited comments decrying the						
1. Publicity accorded ESP work	7	3	11	1	0	22
2. Commercialism of ESP	3	2	0	0	0	5
3. Term ESP, and its definition	14	4	11	0	1	30
. . . . .	. .	. .	. .	. .	. .	. .
Total filled-in questionnaires returned			352			
Failed to return questionnaires			252			
Refused for various reasons to fill in questionnaires			9			
Total			603			

Professor Rhine's book was the most frequently mentioned basis of opinion, having been checked by 164 of those replying. Next most frequently checked was the JOURNAL OF PARAPSYCHOLOGY. Apparently, the supplied list of sources of information was inadequate, for many wrote in others. Thirty psychologists volunteered the fact that their opinions were based upon their own experimental work.

The nearest to a unanimous response was given to the problem of the legitimacy of a scientific attack upon ESP, eighty-nine per cent of

the replies being in the affirmative. Next, in order of unanimity was the classification of ESP within the province of academic psychology, seventy-six per cent of the replies being yes. Of the doubtful and negative replies to the questions, many were accompanied by the comment that it was a proper subject for investigation, but not for the class room.

The other fields suggested as having a bearing upon ESP include chiefly the biological sciences, physics and mathematics. But fourteen psychologists volunteered the opinion that ESP belonged in experimental psychology and "nowhere else."

Thirty-nine per cent of those replying approved, in general, of the experimental approach to ESP as exemplified by the best of the work published in the JOURNAL OF PARAPSYCHOLOGY. In the qualifications to these replies, there was usually expressed approval of the *direction* of the present movement, but doubt as to the scientific value of the work done thus far. Only fourteen per cent felt that the present movement was not in the right direction. The balance left the question unanswered, many of them penning in the comment that they were not familiar with the JOURNAL.

Further comment on this analysis is hardly necessary, since the facts can best be read directly from the table.

Lengthy replies to the final question, "What criticisms have you, or what better methods to suggest?" were made by many. Many criticisms were made in ignorance of the better controlled experimental work. A number of psychologists called attention to errors of technique which had already been corrected. The following points were also made, usually by those who felt that the published results do not demonstrate ESP.

1. There has been unjustifiable selection of data. Records made "when the subject was apparently not feeling just right" have been discarded.

2. The ESP results are statistical artifacts. To quote one of the replies: "It is like betting on a horseless horse race, or like undertaking to create horses out of the act of gambling. A good deal of money could change hands under those circumstances and considerable sporting excitement be experienced, but no colts would be foaled."

3. There exist uncontrolled sensory cues. This was specifically mentioned by only seven, although many others mentioned the need for better control without specifying the possible error.

4. Significant deviations are the result of "recording errors". This argument was mentioned by only four. The operation of illusions or suggestions was mentioned by three.

5. ESP investigators are insufficiently trained, biased, or fraudulent.

The most general constructive suggestion was that the methodology be varied. Many recommended the use of material having more significance, and more emotional value than the ESP cards.

Many psychologists felt that experiments should be designed to identify the *modus operandi* of ESP, through variation of conditions, both environmental and subjective, which might effect the ESP ability.

A QUESTIONNAIRE DESIGNED TO DETERMINE THE ATTITUDES OF PSYCHOLOGISTS TOWARD THE FIELD OF EXTRA-SENSORY PERCEPTION

JAMES C. CRUMBAUGH

ABSTRACT: This paper presents the results of a questionnaire which elicited responses from 100 representative American psychologists. The questionnaire concerned the attitude of the psychologists toward the research in extra-sensory perception, the Journal of Parapsychology, and the Zenith Foundation telepathy program. The results indicate that approximately sixty per cent of the psychologists responding consider the ESP research of value whatever the results, that the majority of psychologists are withholding judgment on the contemporary ESP studies. The answers are somewhat more unfavorable to the Zenith experiments. Over one-fourth of those answering, however, did not know of them. More than one-third did not know of the Journal of Parapsychology.

This study is part of the thesis for the Master of Arts degree in Psychology submitted by James C. Crumbaugh at Southern Methodist University.

FOR THE PURPOSE of discovering the typical attitudes of psychologists toward the hypothesis of extra-sensory perception, and toward the work which has been done in this field, the questionnaire reproduced below was sent to the heads of the psychology departments of 70 universities. These heads were selected by checking every name which appeared as a department head in the 1937 Yearbook of The American Psychological Association. Many persons who were department heads were not thus designated, but the list of names acquired by this method seemed to represent very well every section of the country and every type of institution. The queries purposely were not sent to any institutions which had upon their staffs members whom we knew to have done work, or to have written articles, with either positive or negative indications, in the field under consideration. This procedure was adopted to avoid embarrassing any individual by asking him to comment upon the work of a colleague at his own institution. In at least two cases, however, this situation did occur. In one instance it was an oversight, and in the other the individual had not published his work at the time the forms were sent.

A total of 100 replies were received from 43 different universities. The institutions answering seemed to be as well distributed as to type and location as those to whom the questionnaires were sent. The results are presented by writing in each blank of the accompanying figure the number of individuals checking each statement. The total number of answers given for each question in no case equals the exact number of individuals who replied to the questionnaire, as in each question several persons checked more than one alternative answer. This could be consistently done in most of the questions. However, some of the alternatives which were checked together were not consistent, and all answers, inconsistent or otherwise, have been recorded as they were received. When two inconsistent alternatives were both indicated, this was taken to signify a position somewhere between the two, but as this could not be exactly evaluated, the answers are reported as they were received.

A copy of the questionnaire follows with figures written in the blanks to indicate the number of respondents who checked the statements opposite. Where two answers, inconsistent or otherwise, are given to the same question, both are recorded. Where a question was omitted or changed, nothing is registered.

QUESTIONNAIRE

Please ask each member of your psychology staff to fill out the following questionnaire. Names will be appreciated but they are not necessary. Please note that in some cases more than one topic under each question may be checked if desired. Any additional comment which you care to make may be appended on the back of this sheet.

Very gratefully yours,
JAMES C. CRUMBAUGH,
Southern Methodist University,
Dallas, Texas.

QUESTION I. What do you think of the possibilities of research in the field of extra-sensory perception?

8 1. The field has no value.
38 2. Its value is doubtful.
55 3. There is value, whether results are positive or negative.
6 4. The field will probably open up "new frontiers of the mind."

QUESTION II. What do you think of the contemporary studies in extra-sensory perception at Duke?

3 1. Don't know of them.
3 2. The researches are valid.
54 3. The researches may be valid, though no conclusions can yet be drawn.
4. They are not valid, because of:
31 (1) Sensory cues.
42 (2) Statistical procedure.
6 (3) Dishonesty of experimenters.
28 (4) Other error *(see text for replies)*.

QUESTION III. What do you think of the Zenith Foundation experiments?

29 1. Don't know of them.
28 2. They may be valid, though no conclusions can yet be drawn.
3. They are not valid, because of:
2 (1) Sensory cues.
15 (2) Statistical procedure.
5 (3) Dishonesty of experimenters.
26 (4) Other error *(see text)*.

QUESTION IV. What do you think of the experiments of Gardner Murphy and of John E. Coover in this field?

44 1. Don't know of them.
48 2. They are of scientific value, though no definite conclusions can yet be drawn.
2 3. They are of no value. If not, why? *(See text)*.

QUESTION V. What do you think of the JOURNAL OF PARAPSYCHOLOGY?

35 1. Don't know of it.
3 2. It is a reliable journal, and researches reported in it can be trusted as much as those in any psychological publication.
33 3. It may contribute something of value, though no definite conclusions can yet be drawn.
33 4. Its contributions are of questionable value. *(See text)*.
4 5. It is absurd. If so, why? *(See text)*.

Name (if no objection).....

The following answers were written in the alternative blanks left in the last four questions for discussion.

In Question No. II, alternative answer No. 4, (4), which left a blank for a statement, "other error," 28 persons checked the blank, but only 26 named the error. The 26 replies received are as follows:

1. "Limitations of experimental situation."
2. "Unwillingness to be definite about experimental procedure."
3. "Incompetence as regards experimental method."
4. "Systematic errors of recording."
5. "General carelessness, clerical errors, etc., etc."
6. "Chance proposition."
7. "Unknown errors, very worthwhile to investigate *as errors*."
8. "Other workers don't verify Rhine's results."
9. "Scientific report of results cannot be duplicated by other experimenters."
10. "Bias in conduct."
11. "Large number of imperfections in experimenting, reporting and recording."
12. "A wish which fathers the findings."
13. "No adequate report of procedure."
14. "Failure to give all pertinent data in reports."
15. "General inadequacy of experimental procedure."
16. "Lack of adequate control; Rhine confuses dependent and independent probabilities."
17. "Subject's foreknowledge of all possibilities."

18. "Inadequate controls. How can one control factors which are unknown, or which operate as extra-perceptive factors?"
19. "Clerical errors are reported."
20. "Bias of experimenters."
21. "'Possible mutations not great enough. Too much chance for habit."
22. "Selection of positive cases."
23. "Apparent choice of the wrong probability curve."
24. "Assumption that laws of probability do not apply."
25. "Experimental procedure."
26. "Lack of control."

In Question No. III, alternative answer No. 3, the same choice was given as in the fourth alternative answer in Question No. II. A total of 26 psychologists checked the alternative, "other error," in the third question, but only 23 named the error. Several typical examples of these replies follow:

1. "Motives questionable; no control; commercialization."
2. "Advertising ballyhoo—time elapse since events took place."
3. "Just an advertising stunt—anecdotes are not experiments."
4. "Good as entertainment—otherwise intellectual dishwater."
5. "Even Rhine has deserted this group of studies."

In Question No. IV, alternative answer No. 3, only two individuals answered that the work of Murphy and of Coover was of no value. The statements made by those two follow:

1. "Unknown errors, very worthwhile to investigate *as errors*."
2. "Unscientific."

In the last Question, No. V, alternative answer No. 5, the statement that the JOURNAL OF PARAPSYCHOLOGY was absurd was checked by only 4 individuals. Their remarks follow:

1. "It assumes that telepathy and clairvoyance are proved, and goes on from there."
2. "Only positive results are published."
3. "Non-existence of ESP."
4. "Biased."

Several comments appended upon the backs of the forms may be of interest.

Professor S., a Massachusetts psychologist, very kindly elaborated upon his opinion as follows:

My replies to Question II are of course contradictory as they stand; the asterisks here as elsewhere refer to an in-between position which I hold and which is not covered by your questionnaire. The position: The Duke research is unquestionably valid in a general way. There are far too many facts explicable only in terms of ESP or something akin to it, to support

the numerous contrary arguments *in toto.* At the same time I feel that conclusions cannot be drawn legitimately (beyond the existence of what is called ESP) from the data themselves, without reference to more remote hypotheses seemingly closed to experimentation. In other words, the inevitable conclusions from the data are likely to be too restricted or else to follow some kind of wild goose chase. . . . I see no way out of the impasse; an "act of faith" of some kind is needed; if one is willing to make it the experimental evidence adds little to the picture except in a misleading way. . . . I feel that the research has demonstrated the existence of ESP, but also the impossibility of controlling it through reactive situations. It is apparently *spontaneous.* The fact that it can be produced in the laboratory at all shows that it is extremely widespread in normal operation.

A New York psychologist, Dr. T., wrote the following:

While it is true that one is often accused of being unscientific as soon as he expresses an attitude hospitable to psychic research, I have myself participated in certain observations and experiments and have also had opened up to me ranges of material which make me fairly well assured that our present formulations of laws and principles in psychology and related sciences do not by any means do justice to the whole realm of psychic phenomena. I incline toward the view that future research in psychology, physics, chemistry, physiology, and possibly other related fields—helped out by philosophy—will lead to a great broadening of our present ideas in this particular field. I look forward with more confidence to the contributions of psychology and physics than to what may be learned in some other quarters.

Another New York psychologist, Professor W., wrote:

I do not say there is no possibility of ESP but even so would maintain that there must be some mechanism which is not *extra*-sensory. Doubt whether the artificiality of card and design naming will ever yield results in getting at traits which, if they exist at all, must manifest themselves spontaneously in situations not under any overt control. Coincidences are not and cannot be separated from bona fide clairvoyant experiences, nor does the process of ex post facto recall provide scientific data.

The answers to the first question reveal that psychologists in general believe that there is value in performing research in the field of extra-sensory perception, regardless of the results obtained.

Question II indicates that psychologists are about equally divided between the opinion that the Duke research is not valid and the belief that it may be valid.

The replies to Question III reveal that a large proportion of those who answered do not know of the Zenith Foundation experiments, and that the majority of those who do know of them do not consider them valid. The combination of advertising and scientific research

does not set well with most psychologists. A large minority, nevertheless, are willing to withhold judgment until future developments.

The answers to Question IV bring out the fact that a large proportion of respondents are unaware of the scientific work done in the field of extra-sensory perception and are acquainted only with the more popular literature. Only two of those who did know of the research of Murphy and Coover doubted its scientific value. These two answers indicate that the psychologists involved were in reality entirely unacquainted with the work in question.

Question V reveals that a large proportion of psychologists were not familiar with the JOURNAL OF PARAPSYCHOLOGY. Yet a large number of these individuals, who, since they did not know of the JOURNAL, could not possibly have evaluated most of the research reported therein, seemed very much inclined to doubt the validity of all positive claims in the field. This, coupled with the fact that a large percentage did not know of the work of Coover and of Murphy, indicates that many have criticized without really having had much scientific basis for judgment. The criticisms appear to have been based largely upon the attitudes that have been instilled into the psychologists by their previous professional training, rather than upon a critical scientific analysis. Among those who did not believe the positive claims were valid, a larger number attacked the statistical procedures utilized in the field than alleged an error on any other ground. The taking of this position, regardless of the fact that mathematicians have approved the methods and that empirical control universally supports them, is another indication that much of the criticism has been based on a priori bias rather than on scientific evaluation.

A REVIEW OF RECENT CRITICISMS OF *ESP* RESEARCH

CHARLES E. STUART

ABSTRACT: Within the past year many critical articles concerning ESP research have appeared. A summary of the points raised, particularly with a view to their bearing upon the needs of the ESP investigator, is presented here. The criticisms are classified as attacks upon experimental methods, evaluative procedures, and concepts and implications.

Mr. Stuart is a member of the staff of the Parapsychology Laboratory at Duke University.

THE PAST year has brought forth many dissents from the conclusions of ESP research. Although the volume of critical discussion has been large, when expressions of opinion, speculations, and minor pedantries are deleted the implications for change of procedure on the ESP experimenter are remarkably few. It is the latter fact which necessitates the present article. If several of the articles here considered had met experimental issues and implications squarely there would be no need to reinterpret them. While an attempt will be made herein to "answer" many of the critical charges, it is hoped that this may be done mainly by carrying out the implications of the critics' own argument and conclusions.

In order to give a rough classification to critical points, they will be considered as attacks upon experimental methods, evaluative procedure, and concepts and implications.

CRITICISM OF EXPERIMENTAL METHODS

The most frequently raised criticism of method has been the observation that the commercial ESP cards printed by the Whitman Publishing Company are so warped by the inking process that the symbols may be read, under certain lighting conditions, from the backs of the cards.[1]

[1] This "discovery" has been attributed to a number of individuals. In fact, the warping did not show up in the large proof sheets examined by a number of people (including the present writer). It was noted immediately when the first shipment of cards reached the Parapsychology Laboratory, and reference was made to the condition in the December, 1937, number of this JOURNAL.

An excellent photograph of this fault in the cards is presented by Kennedy (19).

A general discussion of the ESP cards is given elsewhere by Rhine (see ref. 2). The point of concern here is to note what can be or has been done to rule out explanation of the results of ESP tests as due to sensory cues from the cards.

The most obvious method for eliminating such cues is to screen the cards entirely from the sight of the subject. Kennedy reports doing so with one subject and found an immediate drop to chance scoring. He concludes that the sight of the backs of the cards was found to be a necessary condition for the production of high scores (19, p. 149). Others have found this condition not at all necessary; a result directly contrary to Kennedy's is reported by Humphrey and Clark (13). Rhine's summary of all work which met criteria of screening yielded highly significant results (21). The Pearce-Pratt distance work (8), Warner's "test case" (28), Riess' distance GESP (22) are all clear-cut examples of the inadequacy of a sensory cues hypothesis. Many others report significant work with subject and cards wholly separated.

A second criticism of method has been proposed by Kennedy in an unpublished paper (18) and at the A. P. A. round-table discussion. This suggests that testing methods in ESP are particularly liable to errors of recording, and that these errors will be in the direction of any bias on the part of the experimenter. The striking and important differences between the project proposed by Kennedy and the results of his investigation are not noted in the conclusions of the paper. The project as stated is "to give a common-sense explanation, based upon what is already known about human behavior, for the results obtained in card-guessing experiments reported by Dr. J. B. Rhine and others." The speculative case for expecting errors in ESP records is presented convincingly. The experiment set up allowed a number of sources of error and the results were noted. One agent made "practically no errors." The second made errors "of the type calculated to increase the score above chance but their number was not great." The third made errors in a 16 to 1 ratio in favor of increasing the score. The experimentally determined average was a positive deviation of .087 hits per run. The result is *not* significantly positive for the selected subject.

This experimental result is ignored in the discussion and conclusion. The *unreliable* positive deviation is extrapolated to show that it *would be* significant in 10,000 runs, a significance due simply to the fact that the act of extrapolating assumes the original average to be reliable. The conclusions are stated that "these results are not striking but they

do illustrate the possibilities of finding selected agents who may make these unnoticed errors consistently."[2] . . . "Whenever experimental procedures in telepathy and clairvoyance are so devised that unnoticed recording errors are permitted, extra-chance results are to be expected." These conclusions are exactly contrary to the facts that his own experiment, permitting more sources of error than any published ESP experiment, did *not* give extra-chance results, and that the consistency of error by the selected agent was precisely *not* illustrated by the results.

Kennedy reported later work on this topic at the ESP roundtable (2). In this the bearing of the results is unfortunately obscured by the inclusion of records turned in by "an elderly lady," records which gave every indication of an attempt at deception instead of error. Even so, as the graph on page 263 illustrates, the averages are so small that any attempt to use them as a "common-sense explanation" of the compared ESP scores is manifestly ridiculous.

Experiments concerning the Hansen-Lehmann hypothesis of unconscious or involuntary whispering as an alternative to telepathy have been publicized by Kennedy.[3] (1;3) It was shown by Hansen and Lehmann that with the aid of parabolic reflectors subjects could detect involuntary whispering by the agent when he concentrated upon keeping some word or object in mind. (10) In regard to the bearing of these hypotheses, one may point to the absence of parabolic reflectors in ESP experiments, to the fact that the hypothesis has been well known to ESP investigators, and to such evidence as the remarkable Riess series. Unconscious whispering has a place in the history of parapsychology. It cannot be seriously offered in explanation of recent work. It is, of course, completely irrelevant to the large body of ESP research in which the card order is unknown to the experimenter.

[2] It is enlightening to imagine what would be said if an ESP investigator reported the results of tests of three subjects in the form: "One subject chance, one slightly above chance, a third with an average of 5.087 for 184 runs. Since the latter is an experimentally-determined rate for ESP, in 10,000 runs the scores would be highly significant. These not striking results do, however, illustrate the possibility of selecting consistent ESP subjects!"

[3] Kennedy's somewhat catholic methods of presenting his results tends to obscure a research program the real value of which the present writer has no wish to minimize. It is unfortunate that he has committed himself to his conclusions (that ESP results will be "exploded" [3]) before the program is more than under way. His project seems to consist of working out every alternative hypothesis, and to apply it to ESP research as far as it may bear. Ultimate success would consist in pruning down ESP evidence until none would remain incapable of being reasonably explained in more common-sense terms. To do so would be good science. But a straightforward facing of the evidence is just as essential here as in any other research project. This has not been in evidence in Kennedy's reports to date

CRITICISM OF MATHEMATICAL EVALUATION

Much discussion of the mathematical criticisms appearing in recent articles may be saved by examining in outline the data to which mathematical analysis may be applied.[4]

a. The subject call series. The subjects' calls may be studied for preferences of three kinds: (1) total frequency of symbols called, (2) frequency of symbols called with respect to the rank order of the call within the run of 25 calls, and (3) occurrences of specific patterns of calling.

b. The card series. When ESP decks are used, the total frequency of symbols occurring is equal for the five symbols. Their rank order, as well as the occurrence of fixed patterns, may be studied statistically.

c. The correspondences between call and card series. The number, rank order, and patterning of call-card correspondences may be treated statistically.

d. The correspondences between the call series and some other denoted series may also be the object of statistical investigation.

e. Any statistics computed from the above observations may themselves be the object of statistical study.

Statistical studies of subject preferences for symbols have been reported by Willoughby (29) and are illustrated in Table V in the Pearce-Pratt analysis (8, p. 219). The data most frequently expressed statistically in ESP work has been the number of correct correspondences between call and card series. Contingency tables (such as that in the preceding reference) analyse *all* correspondences. Graphical study of rank order of correct correspondences is exemplified in reported DT curves. Studies of the correct correspondences between the call series and series other than the intended card series have been made (8, p. 220).

The application of statistics has been made customarily to the card-call correspondences in an attempt to find whether the number of correct correspondences observed exceeded significantly the number expected upon a "chance" hypothesis. The "chance" hypotheses used have been binomial expectation and matching expectation. This and other methods have been discussed explicitly in this JOURNAL (8; 27).

[4] The Heinleins' statement that "Parapsychologists . . . adopt . . . statistical frames as interpretative indices of *data which cannot be observed* . . . ; the datum of parapsychology is 'extra' to observation . . ." (11, p. 146) reveals the absurdity, not of parapsychological data, but of their attempts to apply a definition from existential psychology to an experimental field. The fundamental "datum of parapsychology" is, of course, the observed response of a subject to an observed situation.

At this date there is little excuse for vague critical generalizations concerning the inappropriateness of the normal curve, the inapplicability of binomial theory, etc.

The mathematical criticisms aimed at ESP research by Dr. C. E. Kellogg have been discussed before in this JOURNAL (26; 27; 15) and repetitive discussion of them here is unnecessary. One of his suggestions, however, that the standard deviation used to evaluate the average score of an experiment should be the standard deviation calculated from the experimental data rather than that fixed by a chance hypothesis, he clarifies further in a recent article (17).

He points out that even when we have worked out the chance hypothesis exactly, the evaluation of the experimental data is still the problem "of the significance of the difference between two mean scores, that by chance and that obtained in the experiment. The formula is altered only by the fact that the chance mean, now exactly known by theory, has no variability, so that the standard error of the difference reduces to the standard error of the experimental mean." (17, p. 384).

As I have suggested before (25, p. 64) the above use of the observed standard deviation is a valid statistical technique, but it assumes the experimental scores to be the best basis for calculating the expected variance. The result one gets by applying this technique is a measure of the significance of the difference between the observed mean and the chance mean on the assumption that the experimental data represents the moments of a chance population from which it is assumed to be a sample. The very test of an ESP hypothesis, however, requires that if a positive result is to be found, it must show up in some non-chance characteristic of the experimental results. The mean is not the only moment of a distribution which may deviate from chance. The use of the theoretical standard deviation still seems the best measure of the expected deviation of an observed average from a theoretical chance mean.

Dr. Kellogg has himself constructed mathematical hypotheses which he holds are more correct than the binomial to express theoretically the chance expectation of a distribution of scores. The only respect in which these various hypotheses differ is in the standard deviation and higher moments. Much controversy has centered about these very small differences. Now all this mathematical work is discarded in favor of the assumption that the observed variance is the best measure of chance variance.

While the context makes it uncertain whether a general statement is

really intended, Kellogg's remark that "the standard error would in all the cases to be evaluated be greater than that for pure chance" deserves the utmost critical inspection. This would certainly be untrue if the "pure chance" hypothesis were an approximation to the experimental results. If it is meant that a widely deviant mean necessitates a greater standard error, the statistical assumption regarding the independence of moments is violated. Indeed, the only basis upon which the statement might be considered true is an empirical one, that distributions of scores of ESP tests do seem to have a greater variance than would be expected. Examples of this are frequent but no comprehensive study is yet available. Such a result would be a very final argument against using the experimental variance to test a chance hypothesis.

In an article entitled "Critiques of the Premises and Statistical Methodology of Parapsychology" (11), C. P. Heinlein and J. H. Heinlein set out to criticize and correct the whole field of experimental work in parapsychology. The authors' attempt to be exhaustively general, and the great number of dyslogistic allusions employed, almost totally obscure the straightforward argument. Considerably abridged, the mathematical thesis is as follows:

Since variously skewed curves of distribution have been shown to occur normally in nature, attempts to force all psychological data into the pattern of the "normal curve" are invalid.

Parapsychologists adopt arbitrarily as "the principal criterion of telepathy . . . the Bernouillian index of abmodality" (called "critical ratio"). "Acts which extend beyond the average 'chance' range of the normal curve of errors and which manifest a mean value skewed to the left or right of the theoretical normal mean value are acts of a telepathic character. . . . Logic dictates that to the extreme variations which are found so freely in the numerous skewed distributions of psychological data may be ascribed the character and function of telepathy. . . ."

The essential contradiction may be seen by noting "that a mean skewed effect is the adopted criterion of telepathy while the unit of measurement for this effect is derived from the non-skewed normal distribution of chance and random sampling. . . . Inasmuch as the standard deviation of a point binomial is a function of" a discontinuous distribution "and inasmuch as the standard deviation of a continuous distribution, when referred to the normal probability integral is a function" of quite different formal character, "the parapsychologists who freely use the theoretical standard deviation of a distribution as a means of computing the likelihood of a variable *should first establish the nor-*

malcy and fitness of the experimental distribution from which the estimates are obtained."

Methods of statistical analysis are suggested *"to ascertain the validity of the Bernouillian standard errors or the Gausian standard error as a means of expressing the mathematical expectation of a given empirical variable x."* This may be accomplished by computing the moments of the experimental distribution, finding the indices of kurtosis and skewness, and testing, by the Lexian ratio or a coefficient of normalcy, whether or not these constants may be considered those of a normal curve.

The actual curve type of the experimental distribution may then be ascertained by the use of Pearson's criteria. A chi-square test of goodness of fit will give the probability that the observed data may be explained upon some hypothetical distribution. The resulting probability is a test of fitness to the hypothesis,—"it cannot serve as a specific criterion of mental telepathy." If when the observations are tested for goodness of fit to a normal curve a significant probability "indicates that the experimental distribution does not satisfy the parameters of the normal curve, hence the theoretical standard deviation is statistically unacceptable and inapplicable."

The above summarized briefly the methodological suggestions of the Heinlein article. It will be noted at once that their applicability to parapsychological experiments rests upon two considerations: (1) whether the assumptions regarding the use of a wholly mathematical criterion in these experiments are true, and (2) whether the technique outlined will give the experimenter information useful to his problem.

Parapsychologists have adopted (arbitrarily) the critical ratio as a measure of performance in ESP tests. But it is obviously not the only criterion of ESP, nor is it even a "principal" criterion. The conditions of the experiment aimed at excluding sensory perception, rational inference, and other alternative hypotheses are equally "criteria" of ESP. The mathematical treatment is merely a test of the extra-chance nature of the performance. Neither logic nor common sense dictates that the mathematical treatment has any implications other than for the ESP test situation.

The common error of confusing two statistical tasks is apparent in the above suggestion. An analysis may be made of the data in order to *describe* it statistically. Fitting a curve to the experimental data does just this. Or a treatment of the data may attempt to evaluate its *difference* from a theoretical expectation, as is done in a chi-square

test of goodness of fit. It has been assumed, and later shown empirically (5) that the binomial distribution is a good fit to chance expectation in these tests. If an extra-chance factor enters into the production of the experimental results there is no *a priori* reason why it should produce a normal, binomial, or any other distribution. We might just as well assume to begin with that the type of the distribution will be unknown. But since we know the type of chance expectation, and measurements *upon the assumption that the experimental results are chance deviants of that expectation* show that assumption to be highly improbable, then we can reasonably infer a non-chance factor. The experimenter may, of course, fit a curve to his experimental results, but information as to whether the results are chance or not must come from comparison with a chance hypothesis either theoretically or empirically based, and can be derived from no mathematical description of the test results.

A recent article (4) by a student of Dr. Heinlein presents an interesting thesis, but one rather baffling to anyone familiar with the recent mathematical writings regarding ESP research. Of the 623, 360, 743, 125, 120 different permutations of a deck of ESP cards, three are chosen for presentation in duplicate tables. The *dependent* probability of each card occurring in a particular rank order in the deck is tabulated. The nature of this dependence is not discussed, but can be illustrated simply. A deck of 25 ESP cards contains five each of five different symbols. If this deck is cut at random, and the upper section of the cut spread out face up, we know at once the total number of cards left in the lower section of the cut, and the number of cards of each symbol left in that section. The lower section now constitutes a modified ESP deck in which the symbols may or may not be equally distributed. The *probability* that a *random cut* of this modified deck will produce a given symbol at the cut is a direct function of the frequency of each symbol within the modified deck. Since we know this frequency from our study of the cards which have been discarded, the probabilities of occurrence of each symbol may be said to be *dependent* upon the constitution of the discarded part of the deck, that is, *dependent upon the knowledge we have gained about the modified deck.*

Although the article purports to show the function of dependent probability in ESP data, no such demonstration is given; possibly because, in spite of the thesis, no reported ESP experiment has ever been conducted under the condition that the subject knows each card as it is guessed. And if Miss Becknell had consulted the reports of ESP

research themselves, rather than limiting her bibliography to the remarks of critics, she might have found that the subject knows nothing about the order of the deck until after the 25 guesses are made.

One may set up, with an ESP deck distribution, as many problems in dependent probability as he wishes. Besides the 600 odd trillion noted there are all the possible variations of each (for example, the probability that the last card is a star when the first and fifteenth cards are known, the others unknown). The crucial question is just how the probabilities found will bear upon the findings in ESP tests. As has been shown (27) when the subject does not know the order of the deck the probability of a correspondence between a call and the occurrence of a card symbol is 1/5; and that dependent probabilities which vary with shuffling affect, not the probability of the correspondence, but the standard deviation of the correspondences.

Miss Becknell points out, as have Willoughby and Kellogg previously, that the subject's call series is not determined as would be the calls of a machine but is probably a function of his own ideology. No reasonable psychologist could disagree with this opinion. But Willoughby has shown that subject preference did not correlate with the hit scores in his experiment; and the contingency studies of the Pearce-Pratt series, evaluated in such a way as to hold preferences constant, show that the preference variation by the subject could not account for the high scores made.

Miss Becknell concludes by observing that "the writer believes that 'chance' in such a series as Rhine employs is entirely unpredictable. . . ." A number of studies not included in Miss Becknell's reference list would suggest that her belief is not shared by all writers. (5; 6; 7; 8; 9; 14; 16; 27).

D. L. Herr (12) after finding the use of binomial expectation in evaluating call-card correspondences "unjustified," suggests that the true distribution is that of a Pearson Type I curve. He then proposes to evaluate ESP data by the use of Tchebyscheff's inequality, a probability evaluation that may be applied to any distribution whatsoever.

The Tchebyscheff criterion assumes, however, that nothing is known about the expected distribution of the results. It requires about 11 per cent of the data to exceed three standard deviations for significance. Such a criterion of significance is sufficient, but is much higher than necessary. As Herr states, the expected distribution is known. There is no need, therefore, to use a criterion based upon a general theorem when the criterion may be based specifically upon the known expected distributions.

A general attack upon the question of selection has been made by Leuba[5] (20, p. 220 ff). He suggests that it is not permissible in calculating anti-chance values: *"To select certain batches of data* made by one subject, as, for instance, the data secured between certain dates, and to use those data as evidence for ESP. . . . *To select certain subjects* out of a much larger number of subjects tested, even though *all* the scores of these selected good subjects are used in the evidence for ESP. . . . *To select certain total investigations* out of all the investigations now going on in this country (and elsewhere) as evidence for ESP. . . . In other words, the large (but unknown) number of ESP runs being made at present, both inside and outside university laboratories, makes it impossible to tell what the chances are for a *purely chance* occurrence of any particular number of hits."

These rules are, unfortunately, not as universal as their statement implies. (If they were, the use of all statistical techniques in science would be quickly halted; as no investigator could prevent someone from repeating his work, destroying the records, and thus introducing an unknown factor into the original investigation.) The anti-chance value of any selection may be calculated in terms of a probability. If the type of selection made is described and the background of data from which it comes is known even roughly, the probability may be evaluated in common-sense terms. Riess' report of his high scoring subject is obviously valid. Any investigator should report all the work of any given experiment. If from that work it is possible to find subjects or conditions which appear to be responsible for the most of the deviations noted, it is wholly permissible to evaluate these separately. Evidence in science can only be found from "selected" investigations. If an investigator sets up an experiment his results may be considered uniquely or as a part of the large number of ESP tests being conducted. Since the mass of work is unknown both as to results and conditions, conclusions based upon a treatment of each investigation as it stands is obviously more reliable.

Leuba's further speculations as to the effect of limiting experiments upon the deviations made by the subject is best answered by such an empirical study as that made by Greenwood (6).

Criticisms of Concepts and Implications

In the glossary, ESP will be found to be defined as "response to an external event (perception) not presented to any known sense." It is

[5] Although evidence from matchings of ESP cards is offered to support his suggestions the exact bearing of the experiments is controversial. For example, the *chance* nature of the matchings is crucial to the conclusions; yet no exclusion of ESP hypotheses was attempted.

obvious that this definition employs concepts that *may* be stated ambiguously, and may lead to any number of implications. As has been explicit in the policy of the JOURNAL OF PARAPSYCHOLOGY, the amplification of the definition of ESP is a task for experimental study. The ultimate aim must be so to qualify each basic concept and implication by experimental procedures and results that ambiguity in the definition is reduced to a minimum.

That the end result of an experiment is a limitation upon the definition of ESP implicit in the procedure is seldom seen in both aspects by the critic. That is, it is possible and necessary to set up arbitrary criteria which the procedure must meet in order that the results may reasonably be considered indicative of extra-sensory perception. But it is also essential to consider whether the arbitrary criteria themselves are necessary to limit a reasonable concept of extra-sensory perception. For example, a correspondent once suggested that he could believe in ESP if a subject would read the text of a book held behind his back. Certainly a successful performance in such a task would indicate some unusual faculty on the part of the subject. But there are many reasonable definitions of ESP that such a procedure would exclude. The construction in experiment of the conditions necessary and sufficient to meet the glossary definition of ESP is not a matter for hasty speculation or dogmatic assertion. And conditions which cannot be interpreted into forms of experimental procedure are of doubtful use in understanding extra-sensory perception.

A general set of conditions to meet certain "scientific" requirements are proposed by the Heinleins.

"Until the parapsychologists demonstrate by the method of concomitant variation a consistently significant point to point correspondence between temporally parallel qualitatively homogeneous experiences (i.e., one experience consisting of an unequivocal conscious recognition of specific telepathic content on the part of the recipient, the other experience represented by an experimenter's voluntary conscious projection of the same specific content not limited by any arbitrary statistical frame and unknown to the recipient before the time of projection) their present extravagant claims may be dismissed on the grounds of being *unscientific*. . . . That some people are better guessers than others has been known since time immemorial, and no amount of mere card guessing will ever serve to demonstrate scientifically the actual transfer of thought from one mind to another." (11, p. 147).

This frank statement might appear to outline a good basis for a con-

cept of telepathy. But experimentally it raises insuperable difficulties. First, there is no way we may know the experience of another person except through behavioral response. Comparison of the qualities of direct experience is therefore impossible. Second, experiences are by their very nature equivocal. Third, even if a subject could be trained to report part of his experience as specifically "telepathic" no reliability could be placed upon the judgment. Fourth, all statistical frames are "arbitrary." Since the most uniform method to show point to point correspondences is that of statistical evaluation these two requirements are contradictory. Finally, the relation between "mere card guessing" and other parapsychological techniques is a problem for experimental study. No references are given to the experimental literature showing that "some people are better guessers than others," knowledge of which would be of the utmost interest to ESP investigators.

Rogosin (23) makes the same error as the Heinleins in assuming that the conclusions of ESP experiments have depended wholly upon the mathematical evaluation. As a result he finds it "crucial" to point out that "the theory of probability cannot *prove* telepathy or clairvoyance a demonstrated fact." He admits, however, that probability theory can test a chance hypothesis concerning experimental results. But, granting this, "there is some question as to whether, given the particular conditions and premises used by Dr. Rhine, it is possible to determine reliably what are the chance figures that should be expected. Secondly, if for the sake of argument it is admitted that his figures are reliably above chance, it is still inadmissible to hypothesize factors which are so diametrically opposed to all scientific knowledge."

In this final proposition, Rogosin explicitly goes a step further than any other critic. He holds that, no matter how scientifically rigorous the experimental methods and results, the hypothesis of ESP is itself inadmissible to scientific inquiry. Before examining his reasons, the direction of the argument may be clarified if we look a moment at the basis of the ESP hypothesis. The problem has not arisen *de novo* in speculative theses or in the laboratory. There is a long history of daily life experiences reported which *seem to have* orderly relations not satisfactorily explained by generally accepted psychological theories. Any individual scientist may *avoid* the problem by noting that the hypothesis of ESP does not arise inferentially from the accepted theories and known facts of science. But this rule of procedure does not *deny the existence* of the problem for science. If it did we should have to agree that all scientific findings must be directly inferable from previous

theories and facts of science. The reverse has actually been true—science has progressed by new theories and facts giving new interpretations and order to previous knowledge.

Rogosin reasons that "Aristotle, Hobbes, and Locke all dealt with this [hypothesis] and rejected it. Modern experimentation has backed them up on the point that knowledge is gained either primarily or indirectly through the senses and cannot be gained without them. Thus, ESP negates centuries of sound work." (24, p. 47). And that ESP is "diametrically opposed to all scientific knowledge. Anything that is based in any way on the idea that entities are not bounded by the limitations of time and space, cannot be accepted as true. Nothing in science today can be claimed to be independent of space or time, and anything which contradicts that may be taken to be non-scientific on its face." (23, p. 269).

As a statement of creed Rogosin's position is sound. He avoids the ESP hypothesis by stressing the speculative work from which it cannot be inferred. But concerning the establishment of the hypothesis itself he has been merely dogmatic. No scientist would accept the writings of Hobbes and Locke as being sounder than modern experimental procedure. And it is obvious that the only way to find out whether knowledge may be gained without the senses is to test the matter experimentally in just such form as ESP experiments have been conducted. Speculative theories concerning the role of space and time in ESP tests as well as in all science are the *result*, not the basis, of experimentation.

* * *

There is no doubt that critical discussions of ESP research will continue to appear in the psychological journals. Their quality should improve. The time seems propitious for a really serious critical survey of the field. If such a survey were graced by a willingness to face the issues directly in terms of the evidence and the experimental methods available the present difference between critical and experimental conclusions might be happily resolved.

REFERENCES

1. ————. Do you think out loud? *Popular Science Monthly*, 1938, **133**, 70-71.
2. ————. The ESP symposium at the A. P. A. J. PARAPSYCHOL. (this number).
3. ————. Unconscious whispering. *Time*, Aug. 8, 1938.
4. BECKNELL, E. A. The function of dependent probability in ESP data. *J. Gen. Psychol.*, 1938, **19**, 373-381.

5. GREENWOOD, J. A. Analysis of a large chance control series of ESP data. J. PARAPSYCHOL., **2**, 138-146.
6. GREENWOOD, J. A. An empirical investigation of some sampling problems. J. PARAPSYCHOL., 1938, **2**, 222-230.
7. GREENWOOD, J. A. Variance of the ESP call series. J. PARAPSYCHOL., 1938, **2**, 60-64.
8. GREENWOOD, J. A. & STUART, C. E. Mathematical techniques used in ESP research. J. PARAPSYCHOL., 1937, **1**, 206-225.
9. GREVILLE, T. N. E. Exact probabilities for the matching hypothesis. J. PARAPSYCHOL., 1938, **2**, 55-59.
10. HANSEN, F. C. C. & LEHMANN, A. Über unwillkürliches Flüstern. *Phil. Stud.*, 1895, **17**, 471-530.
11. HEINLEIN, C. P. & HEINLEIN, J. H. Critique of the premises and statistical methodology of parapsychology. *J. Psychol.*, 1938, **5**, 135-148.
12. HERR, D. L. A mathematical analysis of the experiments in extra-sensory perception. *J. Exper. Psychol.*, 1938, **22**, 491-495.
13. HUMPHREY, B. M. & CLARK, J. A. A comparison of clairvoyant and chance matching. J. PARAPSYCHOL., 1938, **2**, 31-37.
14. HUNTINGTON, E. V. A rating table for card matching experiments. J. PARAPSYCHOL., 1937, **1**, 292-294.
15. KELLOGG, C. E. Letters and notes. J. PARAPSYCHOL., 1938, **2**, 147-148.
16. KELLOGG, C. E. The problems of matching and sampling in the study of extra-sensory perception. *J. Abn. & Soc. Psychol.*, 1937, **32**, 462-479.
17. KELLOGG, C. E. The statistical techniques of ESP. *J. Gen. Psychol.*, 1938, **19**, 383-390.
18. KENNEDY, J. L. Suggestions concerning the nature and production of "extra-sensory" perception. [in manuscript].
19. KENNEDY, J. L. The visual cues from the backs of the ESP cards. *J. Psychol.*, 1938, **6**, 149-153.
20. LEUBA, C. An experiment to test the role of chance in ESP research. J. PARAPSYCHOL., 1938, **2**, 217-221.
21. RHINE, J. B. The question of sensory cues and the evidence. J. PARAPSYCHOL., 1937, **1**, 276-291.
22. RIESS, B. F. A case of high scores in card guessing at a distance. J. PARAPSYCHOL., 1937, **1**, 260-263.
23. ROGOSIN, H. Probability theory and extra-sensory perception. *J. Psychol.*, 1938, **5**, 265-270.
24. ROGOSIN, H. Some implications of extra-sensory perception. *Psychol. League Journ.*, 1938, **2**, 47-49.
25. STUART, C. E. A review of certain proposed hypotheses alternative to extra-sensory perception. *J. Abn. & Soc. Psychol.*, 1938, **33**, 57-70.
26. STUART, C. E. & GREENWOOD, J. A. Letters and notes. J. PARAPSYCHOL., 1938, **2**, 148-149.
27. STUART, C. E. & GREENWOOD, J. A. A review of criticisms of the mathematical evaluation of ESP data. J. PARAPSYCHOL., 1937, **1**, 295-304.
28. WARNER, L. A test case. J. PARAPSYCHOL., 1937, **1**, 234-238.
29. WILLOUGHBY, R. R. Prerequisites for a clairvoyance hypothesis. *J. Appl. Psychol.*, 1935, **19**, 543-550.

RESEARCH NOTES

AN EXPERIMENT IN EXTRA-SENSORY PERCEPTION, WITH SOUNDS AS STIMULI

Hundreds of thousands of trials, using forms on cards as the objects to be guessed, have been carried out during recent years in experimental studies of extra-sensory perception. Other stimuli used have been colored cards, or, as in the Pure Telepathy tests, the mental image of a symbol or form on which the experimenter concentrated. There are several early studies reported (*Proc. S. P. R.*, **3**, 424-452) in which the subject, while in a state of hypnosis, attempted to respond correctly to tactile, gustatory, and olfactory stimuli presented to the experimenter who was some distance away or in another room. There do not appear, however, to be any controlled laboratory studies reported in the literature in which sounds have been used as stimuli in tests of extra-sensory perception.

Studies using sound stimuli appear to be highly desirable and necessary at this time for a number of reasons. First, sounds are fundamental to the normal method of human communication, and the effectiveness of sounds as stimuli in tests of ESP should be determined. Second, there is the possibility that there are systematic variables present in the ESP tests with visual stimuli which have not as yet been discovered and which might be isolated in tests of ESP with sound stimuli or by other variations in procedure. Third, the use of sound stimuli in a controlled experimental study should serve to increase our knowledge of, and broaden the scope of our attack on, the problem of extra-sensory perception.

Such being the case, an exploratory study was undertaken at Bard College in which the stimuli used were sounds; i.e., short recordings of four instruments and the human voice. A majority of the subjects were also given tests with color cards, in order to make possible a comparison between the use of auditory and visual stimuli in ESP processes.

The subjects in this study were all students at Bard College who took part voluntarily when time permitted. None of them received any form of pay or reward for participation in the work. The results of all the trials taken under the conditions to be described are included in the data reported.

Using a Presto Phonograph-Recorder, recordings on acetate discs were made of a violin, clarinet, piano, cello, and singing voice. Each instrument played a different simple melody, ten seconds in length. The recordings were made in separate sections on the record, so that the experimenter could present the stimuli in any desired order. Several records were made from the "master" recording, to insure an adequate supply of standard stimuli throughout the experiment.

The stimuli were presented in random series of twenty-five. In order to have a purely random series, both in order of stimuli and number of each stimulus, the following method was used: Five blank cartridges, numbered from one to five, were placed in a box and thoroughly shaken. One cartridge was then withdrawn blindly and at random by the experimenter, the number noted, the cartridge replaced, and the box again shaken. In this way, twenty-five selections were made. Each number corresponded to one of the five instruments, and thus a randomly ordered series of twenty-five stimuli was obtained. This series differs from the

ESP card series in that *there is not a fixed number of cases;* i.e., there are not necessarily five of each sound stimuli in a series of twenty-five, although such might be the case by chance in any one series. The chance of any one call being a hit is exactly one-fifth, since each call is independent of the others and the stimuli order is determined by chance selection from an infinite source. The number of hits expected by chance in such a series of twenty-five calls is therefore exactly *five.* The expected hit distribution will be binomial.

In all of the tests with sound stimuli, the subjects were located in the Psychology Laboratory on the third floor of the Science Building on the Bard College campus. During March and the first two weeks of April, the experimenter worked with a Presto Phonograph-Recorder in a room located in a building approximately four hundred yards away from the Science Building. The volume control on the machine was kept at the lowest possible level at all times and tests showed that it was difficult to distinguish the sounds correctly two rooms away from the source. The fact that many subjects consistently got chance or below-chance scores, also served as a check on the possibility of actual physical vibrations from the sound source serving as subliminal auditory cues. During the last two weeks of April and the month of May, the phonograph was located in a room approximately one hundred yards away from the Science Building. (The same checks were applicable at this distance as apply above.) The number of subjects calling a series at the same time varied from single individuals to groups of three to seven, and for each set of calls complete records were kept of the prevailing situation.

Before calling the first set, the subject was told what the five sound stimuli were and how the random order of the series had been determined. It was impressed upon him that the number of any one stimulus occurring in a series could vary from zero to twenty-five, although such a series was not at all likely to occur.

The tests were given in the following manner: The experimenter and the subject or group of subjects started their stop-watches together. Five minutes was allowed for the experimenter to go from the laboratory to the room in which the phonograph was located. Then the first sound stimulus was played during the first ten seconds of the sixth minute. The experimenter waited twenty seconds and during this period the subjects recorded their call. The second sound stimulus was then played for ten seconds, the previously determined random order being followed. After a twenty-second pause, the third sound stimulus was played, and so on until a series of twenty-five stimuli had been completed. After an interval of two minutes and fifty seconds, a second set of twenty-five sound stimuli was played, following the same timing arrangement as for the first series; but, of course, in a different random order. Thus, the series of sound stimuli were synchronized with periods during which the subject concentrated on the task and made his call. A period of rest was interposed between the first and second series of stimuli. No more than two series of twenty-five stimuli were given to any one subject during a single test period.

Immediately after the completion of the second series, the experimenter returned to the laboratory with the record of the order in which the sound stimuli had been played. *After covering up the subject's record of his calls,* the experimenter wrote the correct stimulus order in the next column. Then the cover was removed and the correct calls, or "hits", were circled by the experimenter in the presence of the subject or subjects in the group. Since all recording and correcting was checked twice by the experimenter and since an independent record of the calls was made by the subject, it was felt to be highly improbable that any checking or recording errors were not detected under this system.

Fifteen of the twenty subjects who took the sound stimuli tests were also given tests with the color cards. A deck of twenty-five color cards, containing five each of five different colors, identical with those used by Carpenter and Phalen

(J. PARAPSYCHOL., 1, 31-43) was used in these tests. The SDT (screened-down-through) technique was followed throughout the experiment. The deck was shuffled by the experimenter behind a screen and placed face down on the table. The subject, without seeing or touching the cards, then called "down through" the deck while the experimenter recorded the calls. The calls were then covered, and as the subject turned up the cards, *without naming them,* the experimenter recorded the correct order. The calls were then uncovered and the correct calls were circled by the experimenter. The major part of the trials in the color card series were completed during the month of March.

The results of the trials in both the sound stimuli and color card series are given in the following table.

TABLE I

SUMMARY OF SOUND-STIMULI AND CARD-GUESSING TESTS

Subjects	Sound Stimuli						Color Cards					
	Runs	Hits	Aver. per 25	Dev.	S.D.	C.R.	Runs	Hits	Aver. per 25	Dev.	S. D.	C. R.
1. *Br.*	6	20	3.33	—10	4.89	—2.04	20	105	5.25	5	8.94	.56
2. *Bj.*	10	44	4.40	— 6	6.32	— .95	20	102	5.10	2	8.94	.22
3. *Crp.*	8	42	5.25	2	5.65	.35	....	.....	.....	.....	.....	.....
4. *Car.*	10	47	4.70	— 3	6.32	— .47	20	101	5.05	1	8.94	.11
5. *Fr.*	30	169	5.63	19	10.95	1.73	30	166	5.53	16	10.95	1.46
6. *Ga.*	16	82	5.10	2	8.00	.25	.....	.....	.....	.....	.....	.....
7. *Hol.*	2	15	7.50	5	2.82	1.76	30	131	4.36	—19	10.95	—1.73
8. *Hob.*	2	11	5.50	1	2.82	.35	20	106	5.30	6	8.94	.67
9. *Kn.*	2	10	5.00	0	2.82	.....	.....	.....	.....	.....	.....	.....
10. *Me.*	4	25	6.25	5	4.00	1.25	.....	.....	.....	.....	.....	.....
11. *Op.*	30	172	5.73	22	10.95	2.01	45	227	5.04	2	13.41	.15
12. *Rad.*	10	29	2.90	—21	6.32	—3.32	25	153	6.12	28	10.00	2.80
13. *Rue.*	25	140	5.60	15	10.00	1.50	25	134	5.36	9	10.00	.90
14. *Sch.*	28	125	4.47	—15	10.58	—1.41	45	248	5.51	23	13.41	1.71
15. *Sei.*	11	65	5.90	10	6.63	1.51	10	39	3.90	—11	6.32	—1.74
16. *Sto.*	11	51	4.63	— 4	6.63	— .60	25	128	5.12	3	10.00	.30
17. *Ter.*	2	5	2.50	— 5	2.82	—1.76	10	45	4.50	— 5	6.32	— .79
18. *Tra.*	4	16	4.00	— 4	4.00	—1.00	.....	.....	.....	.....	.....	.....
19. *Whi.*	8	47	5.88	7	5.65	1.23	20	98	4.90	— 2	8.94	— .22
20. *Wig.*	4	18	4.50	— 2	4.00	— .50	....	.....	....	....	.....	.....
Totals	223	1,133	5.08	18	29.86	.62	345	1,783	5.17	58	37.14	1.56

From the table, it is evident that the total average for the sound stimuli series, and that for the color card series, are not significantly above chance expectation. To discuss these results as directly indicating anything about extra-sensory perception is therefore highly questionable. Individual features of the results are, however, of interest.

Of the twenty subjects tested with sound stimuli, none scored positively to the extent of a significant critical ratio. One subject gave a significant negative score. Of the fourteen subjects tested with the color cards only one scored significantly high, the same subject who gave a significant negative score on the sound stimuli. A rank correlation of the averages noted for the fourteen subjects tested with both stimuli gives ρ equals —.38. Although neither the correlation nor the averages are significant (with the exception noted above) the trend appears worthy of further investigation. The relatively small number of subjects

and of trials carried out, particularly in the tests with sound stimuli, is another factor which prohibits any attempt at drawing conclusions or even tentatively interpreting the results.

ROBERT SHULMAN.

Bard College.

THOULESS RESULTS

In 1936, Dr. Robert H. Thouless conducted a series of ESP tests to the extent of 6,025 trials, using with the exception of 500 trials the GESP procedure. He obtained a total deviation of —28. Dr. Thouless also compared the empirical score frequency with the theoretical score distribution and found no significant difference. Likewise, the number of sequences of 2, 3, and 4 hits were not found to be other than would be expected by chance. Eleven subjects were involved, with a minimum of 100 trials per subject.

ROTHERA EXPERIMENTS

Under the title of "A Study In Extra-Sensory Perception," Mr. Ralph Edward Rothera submitted a thesis to the faculty of the State Teachers College at Fitchburg, Massachusetts, last June and was awarded the degree of Master of Education.

Rothera's total results are clearly significant statistically. Although they are not pooled for all subjects, the totals are given for each of the eight subjects. The subjects average from 5.27 to 6.18 hits per 25. The average number of runs per subject was approximately 100, a total of 816 runs being made by the eight subjects. The critical ratios for the different subjects range from 1.35 to 6.9, five out of the eight having critical ratios of above 3.

Five different test procedures were used by Rothera, the aim being to give the subject about 20 runs in each of these conditions. The first was ordinary DT, the second OM and third BM without any modifications or additional precautions. The fourth and fifth procedures consisted of GESP—the fourth, GESP as commonly carried out, with the subject ten feet away seated with his back to the experimenter. The experimenter and the cards were protected by a screen, the dimensions of which are not given but which appear from the photograph shown to be adequate to screen all but the upper part of the experimenter's head. The fifth procedure varied in that there is a controlled time interval between trials, an electrical apparatus having been developed for the purpose of signaling when the trial should be given. The signal for a trial was given every ten seconds.

In addition to the tests just described a group test was given to 41 subjects using the GESP test procedure with 10 runs of 25 trials each. In these GESP group tests the sender was located in an adjoining room, signals being given electrically to indicate the trials.

In all of Rothera's work the subject was not informed of the scores until the end of a given series.

The weakness of this work, as Rothera himself recognizes at this stage of it, lies in the limited precautions against possible sensory cues. The best precautions may be regarded as those contained in the group tests in which the sender was in an adjoining room. The results here were low, even considerably on the negative side, giving an average of 4.83 for the 410 runs.

The next most effective safeguarding against sensory cues lay perhaps in the DT procedure particularly since the subject was not seated so as to look at the deck of cards and accordingly could not see even the top card, the only one the back of which would be exposed. The results of the DT test, although they are positive, are not significant. The other four procedures yielded significant results.

A slight difference in average was found in favor of the uncontrolled time interval in the GESP test. The uncontrolled average 5.69 and the controlled time interval 5.59.

The results of the open matching and blind matching tests cannot be taken as evidence of extra-sensory perception in view of the fact that throughout the tests the commercial ESP test cards were used and these are, as was pointed out in these columns, December, 1937, somewhat defectively printed and are subject to identification from the back under certain conditions of lighting. The possibility of recognition of the symbols from the backs of the cards even unconsciously, is sufficient to endanger conclusions from the use of the cards in unscreened tests.

In the GESP tests, which gave the highest averages and which represented the largest number of trials, the card defects could not afford the subject any advantage even through subliminal perception. The Hansen and Lehmann hypothesis of involuntary whispering must, however, be considered since the sender and receiver are in the same room and since the sender looked at the cards. It is true Hansen and Lehmann had to provide large concave reflectors in order to magnify involuntary whispering to the extent of making it a means of communication. In this case on the contrary there was actually a screen immediately in the front of the face of the sender which would have the opposite effect.

It must be recognized, too, that it is not impossible that other means of unconsciously communicating the identity of the symbols to the subject might have been taken by the sender. Such movements as chair-clickings, foot scrapings, clearing of the throat and the like might conceivably serve as involuntary signals if an opportunity for learning were provided. In view of the fact, however, that the subject did not know his scores until the end of the series, it would appear very unlikely that any such signals could become associated with actual hits and thus become reliable as cues.

It is unfortunate that the conditions were not adequate to exclude all possibility of sensory communications and that the conclusions must be, as they are left by Mr. Rothera, indefinite as to the occurrence of extra-sensory perception in the particular tests given. But if the results be given some tentative consideration they would add to the growing impression, that group tests are not as successful in obtaining positive deviations as are individual tests. The suggestion, too, is that the still more personal relations appearing in the GESP procedure are to many subjects at least more favorable for high scoring than are the purely clairvoyant tests.

J. B. R.

GLOSSARY

In order to avoid constant redefining of commonly recurring terms in papers appearing in this journal, the following definitions are submitted for convenient reference. In case of any discrepancy between Glossary and usage in the text of an article the latter should be followed. Words defined elsewhere in the Glossary are italicized in the text of the definitions.

AGENT: In tests for *telepathy*, the person whose mental states are to be apprehended by the *percipient*.

AVERAGE SCORE = AVERAGE: Average number of *hits* per *run*.

BT (BEFORE TOUCHING): The technique in which the top card of the face-down *deck* is *called* and, after being *called*, is laid aside for *checking* at the end of the *run*. Each card in the *deck* is treated in the same way.

BM (BLIND MATCHING): The technique in which the *subject matches* a *deck* of *ESP cards* to five *key cards* which are laid out face down before him in an unknown order. Unless otherwise stated, the order is also unknown to the experimenter.

CALL v.: To designate a card by naming one suit either orally or by indicating it manually (as pointing or writing).

CALL N.: The response described above; also, the resulting selection.

CHANCE: The complex of undefined causal factors irrelevant to the purpose at hand.

CHANCE EXPECTATION = MEAN CHANCE EXPECTATION: The most likely *score* if only *chance* obtains.

CHANCE AVERAGE: *Mean chance expectation* in terms of *average* per *run*.

CHECK: To determine a *score* after the completion of a *run* by comparing the order of the *subject's* calls with the order of cards in the *deck*.

CLAIRVOYANCE: *Extra-sensory perception* of objective events as distinguished from *telepathic* perception (of the mental or subjective events of another person).

CRITICAL RATIO = X VALUE: The observed *deviation* divided by the *standard deviation*.

DECK: Twenty-five *ESP cards*, five of each suit.

DEVIATION: The amount an observed number of *hits* or an *average score* varies from the *mean chance expectation* or *chance average*. A *deviation* may be total (for a series of *runs*) or average (per *run*).

DT (DOWN THROUGH): The technique in which the cards are *called* down through the *deck* before any are removed or *checked*.

EMPIRICAL CONTROL: An experiment which wholly or partially follows the main experiment with the exception that the conditions are designed to exclude the possibility of *ESP*.

ESP (EXTRA-SENSORY PERCEPTION): Response to an external event (perception) not presented to any known sense.

ESP CARDS: *Cards*, each bearing one of the following five symbols: Star, circle, three parallel wavy lines (called "waves"), rectangle, plus.

ESP SYMBOLS: See plate opposite page 1, this journal Vol. I, No. 1, March 1937.

ESP TESTS: A considerable number of techniques come under this heading, conveniently represented by initials, the principal ones being: *BT, DT, PT, GESP, BM, OM, STM*.

EXTRA-CHANCE: *Significantly* different from *chance expectation*.

GESP = GENERAL (OR UNDIFFERENTIATED) EXTRA-SENSORY PERCEPTION: A technique designed to test the occurrence of *extra-sensory perception*, permitting either *telepathy* or *clairvoyance* or both to operate.

KEY CARD: One of the five cards (where there are five suits) against which the cards of the test *deck* (i.e. *target cards*) in the *matching* tests (*OM, BM, TM,* etc.) are *matched.*

MATCHING: A form of *calling* in which a *target card* is placed opposite the *key card* which the *subject* selects to identify it.

OM (OPEN MATCHING): The technique in which the *subject matches* a *deck* of *ESP cards* to five *key cards* which are face up before him.

PERCIPIENT: The person who perceives. In this journal the common use will pertain to *ESP* rather than to sensory perception.

PT (PURE TELEPATHY) = TELEPATHY: The word "pure" emphasizes the exclusion of *clairvoyance.*

RELIABILITY: *Significance.*

RUN: The presentation and *calling* of 25 *ESP cards* or *symbols* in succession.

SCORE: The number of *hits* made in one *run.*

TOTAL SCORE: *Score* of any number of *runs.*

AVERAGE SCORE: *Total score* divided by number of *runs.*

SCREEN: An opaque barrier used between the subject and the card or agent. The main types of screens will be illustrated in this journal on their first introduction in print.

STM (SCREENED TOUCH MATCHING): The *touch matching* technique with an upright *screen* preventing vision by the *subject* of the handling of the cards by the experimenter.

SERIES: Several *runs* that are grouped consecutively or according to some other principle.

SIGNIFICANCE: The unlikelihood that a given event (or series) will on the average occur by chance more often than once in 150 such events (or series). [Arbitrarily taken to mean a *deviation* in the expected direction such that the *critical ratio* is 2.5 times the *standard deviation* (or four times the probable error) or greater.]

STANDARD DEVIATION: A unit of measurement in statistical method; that *deviation* above and below *mean chance expectation* which is expected to include about ⅔ the *chance scores.* For *ESP cards,* S. D. $= 2.04 \sqrt{\text{no. of runs.}}$

SUBJECT: The person who is experimented upon. Most commonly the *percipient* in *ESP,* though also the *agent* in *telepathy.*

TARGET CARD: The card which the *percipient* is attempting to perceive (i.e. to *call,* or otherwise indicate a knowledge of).

TARGET DECK: The *deck* of cards the order of which the *subject* is *calling.*

TELEPATHY: *Extra-sensory perception* of the mental activities of another person. It does not include the *clairvoyant* perception of objective events.

TOUCH MATCHING: The technique in which the *subject* indicates his *call* by tapping or touching one of the five *key cards* while the *experimenter* places the *target card* thus called in front of the *key card* indicated.

TRIAL: A single *call* of a presented *symbol.*

X VALUE = CRITICAL RATIO: The observed *deviation* divided by the *standard deviation* or by the *probable error.*

Volume 2 DECEMBER, 1938 Number 4

The Journal of Parapsychology

WILLIAM McDOUGALL *and* JOSEPH BANKS RHINE, *Editors*
CHARLES E. STUART, *Assistant Editor*

Contents

Page

William McDougall (1871-1938) 239

Editorial Comments 242

The *ESP* Symposium at the A. P. A. 247

A Comparison of Blind and Seeing Subjects in *ESP* Tests . . . 273
MARGARET M. PRICE

Studies in Extra-Sensory Perception: II. An Analysis of a Second Series of 25,000 Trials 287
DOROTHY R. MARTIN and FRANCES P. STRIBIC

A Survey of Psychological Opinion on *ESP* 296
LUCIEN WARNER and C. C CLARK

A Questionnaire Designed to Determine the Attitudes of Psychologists toward the Field of Extra-Sensory Perception . 302
JAMES C. CRUMBAUGH

A Review of Recent Criticisms of *ESP* Research 308
CHARLES E. STUART

* * *

Research Notes 322

Glossary 327

DUKE UNIVERSITY PRESS
DURHAM, N. C.

The Journal *of* Parapsychology

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

This journal is published on the first day of March, June, September, and December by Duke University Press, Durham, North Carolina.

Contributions submitted for publication and all editorial communications should be addressed to the Assistant Editor, Mr. Charles E. Stuart, College Station, Durham, North Carolina. Correspondence with the editors is advised before submitting articles other than reports of experimentation. Since it does not forward manuscripts by registered mail, Duke University Press 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 THE JOURNAL.

Reprints should be ordered when the proof is returned. Fifty copies are supplied to the author free; additional copies at cost. Correspondence concerning subscriptions, change of address, back numbers, and other business communications should be addressed to Duke University Press.

The subscription price is $3.00 a year; single current numbers 75c. The rate for back volumes is $4.00; for single numbers $1.00. Missing numbers will be supplied free when lost in the mails if written notice is given within one month of the date of issue. All remittances should be made payable to Duke University Press.

Entered as second-class matter April 19, 1937, at the post office at Durham, North Carolina, under the Act of March 3, 1879.

COPYRIGHT, 1938

BY

DUKE UNIVERSITY PRESS

The Journal *of* Parapsychology

Contents of Preceding Issues of Volume 2

MARCH, 1938

Editorial Comments

The Case of Ilga K.: Report of a Phenomenon of Unusual Perception
HANS BENDER

Studies in Extra-Sensory Perception: I. An Analysis of 25,000 Trials
DOROTHY R. MARTIN and FRANCES P. STRIBIC

A Comparison of Clairvoyant and Chance Matching
BETTY M. HUMPHREY and JOHN A. CLARK

Experiments Bearing on the Precognition Hypothesis
J. B. RHINE

Exact Probabilities for the Matching Hypothesis
T. N. E. GREVILLE

Variance of the *ESP* Call Series
J. A. GREENWOOD

JUNE, 1938

Editorial Comments

The Experimenter-Subject Relationship in Tests for *ESP*
J. G. PRATT and MARGARET M. PRICE

A Study of Card-Guessing in Psychotic Subjects
ROBERT SHULMAN

The Tyrrell Apparatus for Testing Extra-Sensory Perception
G. N. M. TYRRELL

Experiments Bearing on the Precognition Hypothesis: II. The Role of *ESP* in the Shuffling of Cards
J. B. RHINE, BURKE M. SMITH, and J. L. WOODRUFF

On the Combination of Critical Ratios
FRANK W. BUBB

Analysis of a Large Chance Control Series of *ESP* Data
J. A. GREENWOOD

SEPTEMBER, 1938

Editorial Comments

Discrimination Shown Between Experimenters by Subjects
JAMES D. MACFARLAND

The Effect of Rate of Movement in Card Matching Tests of Extra-Sensory Perception
CHARLES E. STUART

An Unusual Case of Extra-Sensory Perception
RALEIGH M. DRAKE

ESP Tests With Enclosed Cards
J. B. RHINE

An Experiment to Test the Role of Chance in *ESP* Research
CLARENCE LEUBA

An Empirical Investigation of Some Sampling Problems
J. A. GREENWOOD

PERIODICALS PUBLISHED BY DUKE UNIVERSITY

American Literature. A quarterly journal devoted to research in American Literature, published with the coöperation of the American Literature Group of the Modern Language Association of America. Subscription, $4.00 per year. Back volumes, $5.00 each.

Character and Personality. A psychological quarterly devoted to studies of behavior and personality. Subscription, $2.00 per year. The first number was published September, 1932.

Contributions to Psychological Theory. A monograph series dealing with problems of psychological theory in the widest sense, including their relations to other fields of inquiry. The monographs appear irregularly. Subscription, $5.00 per volume of approximately 450 pages.

Duke Mathematical Journal. Began publication in 1935. Published under the auspices of Duke University with the coöperation of the Mathematical Association of America. Subscription, $4.00 a year.

Ecological Monographs. A quarterly journal devoted to the publication of original researches of ecological interest from the entire field of biological science. Subscription, $6.00 per year. The first number was published January, 1931.

Educational Monographs. $1.00 a number.

Hispanic American Historical Review. A quarterly review dealing with the history of the Hispanic American countries. Subscription, $4.00 per year.

Historical Papers of the Trinity College Historical Society. Series I-XX previously published. $1.00 a series as available.

The Journal of Parapsychology. A scientific quarterly dealing with telepathy, clairvoyance, and other parapsychological problems. Subscription, $3.00 a year.

Law and Contemporary Problems. A quarterly published by the School of Law, presenting in each issue a symposium on a problem of current importance having significant legal aspects. Subscription, $2.00 per year. The first number was published September, 1933.

The South Atlantic Quarterly. A magazine of modern opinion and discussion, founded in 1902. Subscription, $3.00 per year.

DUKE UNIVERSITY PRESS
DURHAM, NORTH CAROLINA