The Journal *of*
Parapsychology

Volume 1 · MARCH, 1937 · Number 1

EDITORIAL INTRODUCTION

SOME FOURTEEN YEARS ago the late Harry Houdini conspired with one of the editors of this journal to bring psychical research to the notice of the universities and to secure for it a place among their recognized fields of study. The first step was to secure the willing and effective co-operation of Professor Carl Murchison, head of the Department of Psychology at Clark University, which university Stanley Hall had made, many years earlier, the headquarters of American psychology. The first product of that conspiracy was a series of lectures at Clark University in the year 1926, later printed together in the volume timidly entitled *The Case for and against Psychical Belief*.[1] In that volume (which has enjoyed considerable circulation) the conspiratorial purpose was openly confessed in an essay entitled "Psychical Research as a University Study." In stating the thesis that psychical research should be frankly and fully admitted into the universities, the following considerations were stressed:

The supreme importance of the questions on which it might be hoped that psychical research would throw some light and the universal interest in such questions of intelligent persons of all times and places: What are the relations of mind to matter? Are mental processes always and everywhere intimately and utterly dependent upon material or physical organizations? Do the volitions, the strivings, the desires, the joys and sorrows, the judgments and beliefs of men make any difference to the historical course of the events of our world, as the mass of men at all times have believed? Or does the truth lie with

[1] The lecturers and co-authors were: Sir Oliver Lodge, Sir Arthur Conan Doyle, Frederick Bligh Bond, L. R. G. Crandon, Mary Austin, Margaret Deland, William McDougall, Hans Driesch, W. F. Prince, F. C. S. Schiller, John E. Coover, Gardner Murphy, Joseph Jastrow, and Harry Houdini. The volume was edited by Dr. Carl Murchison and published by the Clark University Press in 1927.

those few philosophers and scientists who, with or without some more or less plausible theory in support of their view, confidently reject well-nigh universal beliefs, telling us that the physical is co-extensive with the mental and that the powers and potentialities of mind may be defined by the laws of the physical sciences?

These and many closely allied questions have vexed the minds of men since men first began to think. Beyond defining more clearly such questions and the various possible answers, the labors of metaphysical philosophers have made no progress with them. The chief result of more than two thousand years of effort along that line has been to array all competent thinkers in more numerous and more sharply divided schools.

Nor until very recent years has science been any more successful. During the modern period, from the time of Galileo to the end of the nineteenth century, science became increasingly inclined to regard mechanical principles as adequate to the interpretation of all events, those of the realm of living things not excepted: and the inevitable implication was that all we call mind or mental activity is but a functionless by-product of increasingly complex mechanical systems. Yet this conclusion, although confidently proclaimed by some, appeared intolerable to many men of science, as well as to most philosophers and intelligent laymen; intolerable not merely because inconsistent with all moral aspiration and religious belief, but also because intellectually disreputable: for this conclusion implied an account of the universe which made of it a closed mechanical system; an account, therefore, which could find no place and no function within the universe for mind; an account which omits all reference to the very mental activities which, with immense effort and marvellous ingenuity, had gradually built up that account, or, if making mention of them, inevitably leaves them virtually outside the system hanging loosely upon it without intelligible relation to it.

In face of this intolerable situation a common practice was to cultivate the compartmental mind, to keep science and its conclusions in one compartment, religion, morals, and practical wisdom in all human affairs in another. Yet such practice was seldom wholly successful in bringing peace of mind, and could never be more than a *pis aller*, an unsatisfactory compromise of unstable equilibrium.

About sixty years ago a few men much concerned with these large questions discerned the truth that men seldom observe or discover what they are not explicitly seeking; they initiated, therefore, a systematic search for evidences of the relative independence of mind and its func-

tions, evidences which science might well have overlooked just because, being inconsistent with its dominant trend to a mechanical view of the universe and all that is therein, they had never been systematically sought for and recorded. The expression, "psychical research," is now generally understood to mean such systematic search. The English Society of that name (which is responsible for the choice of this illogical and unsatisfactory designation) emulated by similar groups or societies in other countries, has prosecuted such research for more than fifty years, and not without encouraging successes. The work of these societies has been almost entirely the hobby, the avocation, the leisure-hour pursuit of persons who (though in many cases highly distinguished in various other fields of endeavor) have remained amateurs in this field. Now there is no room for doubt that this enterprise is more difficult than any other branch of empirical science. If, then, the other branches of science have thriven and built up bodies of well-founded knowledge and illuminating theory, only by means of the co-operative endeavors of many highly trained specialists working in laboratories and observatories, with specialized methods and instruments provided by universities and other great institutions of the higher learning, what hope of similar success in this most difficult branch so long as it is prosecuted without similar advantages? No wonder that sixty years of endeavor of this amateur kind, very admirable as it has been in many ways, has not succeeded in gathering evidences which would suffice to establish for men at large definite answers to the questions asked! These efforts have succeeded in effectively raising the question of the possibility of attaining, by the methods of science, answers to some of the great questions previously left to the metaphysicians; and they have made such inquiries respectable, which in itself was no small achievement. But it has become clear, nevertheless, that most admirable and indispensable as was the work of these societies, something more is required to carry further the advance they have so well begun, work of a more systematic, more sustained, more continuous and more broadly based kind; in short, organized professional effort the continuity of which should not be in immediate dependence on popular favor and interest and on the lives, health, and self-sacrificing labors of a few amateur enthusiasts.[2]

[2] In the year 1911 one of us devoted a large volume *(Body and Mind)* to a critical presentation and evaluation of all the arguments for and against the naturalistic view that mind, always and everywhere, is closely bound up with physical organizations. Although he endeavored to give full weight to the evidences accumulated to that day by psychical research, in support of the *contra* argument, he found it impossible to make that argument such as would compel the assent of all competent and impartial inquiries. The contributions made by psychi-

An endowed research institute at once suggests itself, and at least one such has come into existence and has done good work.[3] But detached research institutions have many intrinsic drawbacks; especially, the lack of opportunities for easy intercourse with workers in other fields, a drawback which in the case of an institute devoted to psychical research must be especially serious. More than any other research worker, the student of psychical research needs to live and work in a community of students and specialists along many lines such as can only be found in the university. He needs contact with them to keep him sane, balanced, critical, yet devoted to the pursuit of problems that may at times seem insoluble, by methods which at times may seem trivial, boring and absurdly inadequate to the immense significance of the problems pursued. He needs to live and work in an atmosphere of many-sided research, critical and skeptical, yet tolerant, understanding, open-minded to all possibilities, one which has achieved some comprehension of the extent of our ignorance and the limitations of our knowledge, one which recognizes that all things are possible save the logically contradictory.

And a special research institution must involve a further drawback; namely, it is not likely to proliferate, to propagate itself by instigating similar activities elsewhere. Whereas, when research of a new type is undertaken in one university with some promise of success, it is likely to be taken up by the corresponding departments in others.[4] And for psychical research such parallel research in a number of relatively independent centers is peculiarly important.

The research under discussion will richly repay the benefits gained from the atmosphere and facilities of a university. For it is one of those frontier studies which have always been fertile in new knowledge and new understanding, and which, contrasting with the ever more detailed explorations of long-occupied territories, bring the stimulating sense of adventurous advance into new regions of limitless possibility. Also, such research demands from its devotees the exercise, and therefore the development in the highest degree, of all the intellectual virtues;

---

cal research since that date, though interesting and important, do not seem to have changed that state of affairs. The same must be said of all the more recent evidences and arguments contributed by philosophy and the sciences. In short, a deadlock has prevailed for sometime; and it would seem that it is likely to continue unless we can apply some more powerful explosive than any hitherto available.

[3] *Institut Métapsychique International,* Paris.

[4] An anticipation that seems to be in process of realization in the present instance; as illustrated by the contents of this first number of our journal.

it requires constant severe self-criticism, a disinterested pursuit of truth, and readiness in such pursuit to face incredulity rooted in massive traditional prejudices, scorn, ridicule, jeers, and loss of reputation; it involves facing most subtle problems of logic and scientific method, and grave risks of self-deception in spite of the best intentions. It requires, in short, such severe discipline as may lift the devoted worker almost to the level of Robert Browning's grammarian.

And in another way the university which gives countenance and support to research of this kind may find that it has not in vain cast its bread upon the waters. There are many friends and supporters of higher education who have come to doubt the need for ever more hurried pressing forward with research in the physical sciences, persons to whom it seems that already the advance of physical science and the multitude of its practical applications has thrown the world into chaotic condition from which recovery is problematical and to which physical science can contribute nothing but further unbalance and disorder of our social life. Some such persons may have come to think that increase of material wealth does not always necessarily mean greater welfare, that increase of complication of the mechanical bases of our civilization is not necessarily the most desirable form of progress; they may even incline to think that we may reap more advantage from increased understanding of man, his nature and possibilities, than from increased intimacy with the remoter nebulae or from the meteoric appearance of yet another theory of the atom.

Yet another consideration was urged; namely, the responsibility of the universities for the guidance of the public, for helping the popular mind to distinguish between truth and error in a sphere in which the drawing of the distinction is peculiarly difficult, to defend itself against the ill-founded claims to new and profound truths put forth in an ever-increasing flood by individuals and societies who, in some cases sincerely, in others with mixed and questionable motives, but in general without any special competence or sources of information, manage to secure a total following which at any moment must be numbered in tens of millions.

At the date of the conspiracy here mentioned, the project it sought to promote was not an absolute novelty. Some few steps in the desired direction had already been made. Stanford University had accepted (in the year 1912) a very considerable sum of money bequeathed to it for the promotion of such studies. And smaller donations for similar purposes had been accepted by Harvard and the University of Pennsylvania. But only one of these foundations (namely, the one at Har-

vard) had produced (in the early nineteen-twenties) some small positive results. In one European university, namely, that of Groningen, members of the department of psychology had, at about the same date, reported a careful laboratory study of a single subject who seemed to manifest some striking telepathic ability. It may therefore be claimed that when, about the year 1930, Duke University encouraged Dr. J. B. Rhine, a member of its department of psychology, to follow up his already keen interest in this field, it was playing a pioneer rôle and giving a lead to the academic world along a path which few had trodden, a path of uncertain promise and one beset with risks and dangers.

But already it seems clear that the time was ripe for the realization of the aims of the conspirators. Whatever the ultimate outcome of this movement to acclimatize psychical research in the universities, the movement up to the present moment has prospered undeniably. It has proved possible and practicable to promote intensive research in the laboratory of the psychological department, enlisting the keen assistance of a number of senior students without unduly interfering with their academic careers or in any way disrupting the work of the department. The growth of the work in scope and personnel is by natural and gradual steps leading to the formation of a sub-department. This healthy growth is no doubt mainly due to the fact that the intensive application of the strictest laboratory methods on a scale never previously attempted has produced results which seem to be positive indications of an extension of our knowledge and to give promise of still further extension by the pushing of similar strictly controlled experimental methods. Although it is still too early to speak with any finality of these methods and their positive results, it can at least be claimed that they have excited a very widespread interest and have had great effects in the way of stirring up other scientifically equipped students to repeat and vary the procedures.

The volume *Extra-Sensory Perception,*[5] in which Dr. Rhine published an account of experiments that seem to establish the achievement of both telepathic communication and clairvoyance by a number of selected subjects has been very favorably received both by the press in general and by many men of science. It is perhaps not too much to say that, for the first time, the scientific world has been compelled to take seriously the possibility of modes of perception utterly unlike those

[5] First published in 1934, by the Boston Society for Psychic Research; English edition by Faber and Faber in 1935; second American printing by Bruce Humphries Inc. of Boston in 1936.

hitherto generally recognized, and utterly incompatible with any world-view of the kind accepted by the great majority of men of science.

And these studies in extra-sensory perception have had this large appeal not merely because they provide more convincing evidence of such modes of perception than any previously made, nor yet because they conform to the strictest models of laboratory procedure, but also because they seem to be throwing some light upon the *modus operandi* and to promise more.

Hitherto we have spoken of "psychical research," the name adopted sixty years ago by the founders of the English Society of that name to designate the inquiries of the type they had in view. The title of this journal requires a few words of explanation as to how we conceive the relation of parapsychology to psychical research. Parapsychology is a word that comes to us from Germany, where for some dozen years past it has been used to denote the stricter inquiries into obscure and questionable forms of mental activity. We think it may well be adopted into the English language to designate the more strictly experimental part of the whole field implied by psychical research as now pretty generally understood. It is these strictly laboratory studies which most need the atmosphere and conditions to be found only in the universities; and it is these which the universities can most properly promote, leaving the extra-academic groups the still important task of collecting and recording all such reports of phenomena apparently expressive of unusual mental powers as occur spontaneously, obscure warnings and premonitions, veridical phantasms of the living and the dead, and other sporadic manifestations of mysterious origins.

We do not claim that any sharp line can be drawn marking off the field of parapsychology within the larger vaguer province of psychical research. Rather, we anticipate that the stricter experimental methods will gradually invade other parts of the province annexing them to their own more special field, until possibly the two shall coincide. But we regard the differentiation of the two terms as useful at the present time; and it is our intention to admit to this journal only contributions that properly fall within the narrower sphere implied by its title; that is to say, reports of experimental studies in the stricter sense and discussions of methods and interpretations of such work.

At present it seems likely that there will be sufficient material of the kind indicated to fill four numbers of this journal each year. If it should prove that we are overestimating the supply, we shall not hesitate to postpone publication until we shall have sufficient material of

high quality to make up a number. For the immediate future it seems likely that the supply will outrun our demand. Since the publication in 1934 of the first full report of the work done on extra-sensory perception in the psychological department of Duke University, there has been a considerable extension of active interest in parapsychological research. Repetition with and without variations have been undertaken in more than a dozen universities; and, in addition, several extra-academic groups have reported experiments that seem deserving of a place in our pages. The need for multiple repetitions by different observers of all experiments yielding positive results is far greater in this field than in others, and for two reasons: First, to provide against the peculiarly great risks of *bona fide* error and self-deception; secondly, to carry conviction of the objectivity of reports and validity of conclusions to the naturally and properly skeptical public which this journal seeks to serve. In the absence of some such journal as this, there would be much inevitable delay in publication of practically all contributions of the kind which we invite. All the regular psychological journals seem to have difficulty in finding sufficient space for suitable matter offered to them.

If, then, any of our readers should have trustworthy information concerning what may seem to be some sporadic display of mysterious powers or any other such phenomenon, intriguing but not susceptible of strictly experimental investigation, we shall be glad to hear of it; but we cannot undertake to publish any report of it. We beg him to send any such report to the Boston Society for Psychic Research, or to the London Society for Psychical Research. And, if he is a member of one or both of these societies, we beg that he will not withdraw his support from them on subscribing to this journal. For the publications of those societies will keep him in touch with a wider field of investigation and report than we propose to cover. In this connection we announce that for the near future we do not propose to publish reviews of books; though we shall try to bring to the notice of our readers reports of any work of high significance in our special field that may be published in any part of the world.

We shall do our best to maintain a high critical standard, to guard against errors of all kinds, and to protect our readers against lack of good faith on the part of any contributor. But we shall not pretend to guarantee the accuracy of reports which we hope to publish, except those which may appear over our own signatures.

We are acutely aware of the heavy responsibility towards the public which we are undertaking. It is largely for this reason that we have

abstained from asking the co-operation of others whose names listed as an editorial board might have adorned and given distinction to our cover. The reception by our scientific colleagues of the reports of the parapsychological work done in Duke University has been most generous and gratifying; we can ask for no better reception for this journal. We invite their frankest and severest criticism, believing that such criticism is much needed in this branch, as in all other branches, of psychology.

It is not our intention to restrict this journal to reports on extra-sensory perception alone. But much of the material on hand deals with that special field; and it seems likely to be the principal field of experimental research in the near future. We would therefore express the hope that readers of this journal may have made acquaintance with Dr. Rhine's book *Extra-Sensory Perception;* since, if we may assume such acquaintance, such minimum of common background, our task and that of our contributors will be considerably simplified. However, in order that the reader may have at hand some introduction to this background we reprint, in an appendix (p. 70), a somewhat abbreviated report of some of these basic experiments.

CLAIRVOYANT BLIND MATCHING

J. G. PRATT

ABSTRACT: A subject, one of 124 tested, was found capable of consistent ability in extra-sensory perception. Of 85,000 trials made with this subject, 7,800 are selected and reported as meeting rigorously the criteria for the exclusion of all possibilities other than clairvoyance. These trials were made in the following manner: Five key cards were placed in a row before the subject, who indicated her call by pointing to one of them. The cards she was calling were held face-down behind a screen by the experimenter, and were laid, still face-down, opposite the key card indicated. The key cards themselves, however, were face-down and covered with blank cards, their order unknown to subject or experimenter. The subject scored successfully with this procedure during the period reported. Later, when her ability declined on other tests, her performance in this procedure also dropped to chance. An empirical control affirms the reliability of the scores made.

Dr. Pratt conducted his research in the laboratory of Professor Gardner Murphy of the Department of Psychology of Columbia University.

IN OCTOBER, 1935, the writer began a program of research in extra-sensory perception in collaboration with Dr. Gardner Murphy of Columbia University. The program was planned on the strength of the report of the Duke University experiments, and Dr. Rhine's general methods of investigation were followed. We had in mind two main objectives in undertaking work in this field. First, could results similar to those of the Duke experiments be obtained? This question concerned verification, not in the strictest sense of repeating the experiments under identical conditions, but in the sense of adequately controlling against sensory perception. Second, what new light could be thrown upon the psychological nature of the process (or processes) involved in "extra-sensory perception"?

At the time of writing, the program has been under way for 15 months, and we feel that progress has been made toward achieving both objectives. The purpose of this paper, however, is to report the results only insofar as they bear upon the question of verification. Has this program produced any evidence that knowledge may be acquired without the use of the known senses?

The first five months were largely unsuccessful as far as finding good scorers was concerned. In the preliminary tests made, as in almost all later ones, the material to be "guessed" was a deck of 25 or more cards (ESP cards) made up of sets of 5 symbols (cross, circle, rectangle, star, wavy lines). When a new subject was tried out, he was told a bit about the nature of the experiments and of reports of earlier success. The tests were then given in one of a number of ways, an attempt being made to have the conditions seem as natural and easy as possible. The aim of the preliminary tests was purely to discover good scorers; these first results, no matter how high, could not be considered evidence for the occurrence of ESP (extra-sensory perception). During the first 5 months of preliminary testing, 123 people were tried. In all, they made over 150,000 calls. However, only three of these subjects showed any promise of good scores; and of these three, one was not available for further work, while the other two did not continue to score above chance.

The method of preliminary testing which was successful with these three subjects (and with the fourth one discovered later, who alone has been available for an extended period of investigation) was "open matching." This procedure may be described as a special game of solitaire played with a special deck of cards. One of each of the 5 symbols is placed face up before the subject. The subject then takes the shuffled deck of ESP cards in his hands and lays each card out, face down, opposite the one he "feels" it matches. In a deck of 25 cards, the average chance number correctly placed should approximate 5. When a subject's score deviates significantly from this average,[1] more rigorous conditions, such as the experimenter's holding the deck out of the subject's sight behind a screen, are introduced as fast as the subject can successfully adapt to the new procedure.

Sometimes, however, the subjects quickly lose the capacity to score, or advance to more strict conditions very slowly. Thus the two subjects who were successful in the preliminary tests and who were willing to proceed with the tests made an average of slightly over 11 hits per

[1] The importance of the result rests with the deviation of the subject's total score from the theoretical chance score, one-fifth the total number of trials. The standard error of the deviation is given by the formula $\sqrt{npq}$, where n is the number of trials, p is the probability of success on each one (1/5), and q is the probability of failure on each one (4/5). The critical ratio, or X, of the result is found by dividing the standard error into the deviation, and this is interpreted by the use of the proper tables into the odds against the result being one of mere chance coincidence. An X of 2.5, giving odds of 150 to 1 against chance, is for most purposes considered sufficient to rule out the hypothesis of chance coincidence.

25 cards for the first 750 trials each. Something more than chance coincidence was clearly indicated. But their success ended almost as abruptly as it had begun, and they were never able to score consistently high again under any of a number of conditions tried. These cases are not offered as evidence of ESP, but for the sake of completeness in reporting the events leading to the later discovery of a good subject.

The subject with whom this report is chiefly concerned will be identified as Mrs. M. She was first tested for scoring ability at the open matching procedure on March 22, 1936. The first few thousand trials did not show consistently high results, but occasional series of 250 or 500 trials averaged 7 or more per 25 cards. On the strength of these, the work was continued with the hope that the scores would become stable at an extra-chance level. This hope was realized after a month of preliminary work, and then the subject was introduced to a procedure called "screened touch matching." This new procedure was one to which the subject easily adapted. It was not as rigorous as some others the experimenter had tried during the period of the development of the subject's ability, but it served to facilitate study of the manner in which the scores varied with various changes of conditions.

As with the open matching procedure, the results obtained from the original screened touch matching will not be presented as evidence that ESP occurs. In this procedure a wooden screen, two feet square, stood upright on the table, separating the subject and experimenter. The lower edge of this screen was raised three inches above the table. The row of key cards against which the cards of the shuffled deck were to be matched were placed under the edge of the screen. When seated at the table ready to work, the experimenter and subject were out of each other's sight. During the matching process, the experimenter held the shuffled deck face down in one hand while with the other hand he placed the cards out into five piles as directed by the subject's touching one after another of the key cards. In this way the subject had no physical contact with the cards she was matching.

This original procedure does not, however, completely safeguard against all modes of sensory perception of the order of cards. This is because of the fact that there was an aperture left to allow the pointer to be seen by the experimenter. This aperture might have permitted vision of the cards if the subject brought her head to the level of the table. It is true, the experimenter aimed so to hold the cards as to render them invisible even through the aperture, and usually the pointing was in advance of the removal of the top card from the pack.

Nevertheless, this possibility of vision is enough to warrant rejection of this condition as adequate to exclude sensory perception.

The total number of trials made at screened touch matching gave very significant extra-chance scores. These are very probably genuine instances of knowledge not obtained through the known sensory channels, but this cannot be absolutely concluded from the conditions. The importance of these results lies in their bearing upon the question of how different conditions affect scoring in ESP tests. There are definite indications along this line which will be later submitted following further verification and completion.

The results reported here as evidence of extra-sensory perception are based upon a modification of the screened touch matching procedure, intended further to safeguard the test against possible sensory cues. But in submitting these selected results the reader may wish to have the totals for the screened touch matching as well as those of the more advanced condition. For comparative purposes, therefore, all the trials made under the screened touch matching procedure are presented here. The work falls naturally into four periods, though the division between the third and fourth periods is marked only by the end of extra-chance scoring.

Period	*Trials*	*Per cent correct*
May 5-May 25	17,500	29.6
Aug. 10-Sept. 19	11,650	25.7
Oct. 2-Nov. 13	28,050	25.1
Nov. 14-Dec. 19	28,775	20.0

A general decline will be observed,[2] the first period giving the highest rate of scoring and the last producing only chance. At the time of writing this paper, Mrs. M. is still taking occasional tests, but has not shown a return to extra-chance scoring.

It was during the third period of working, October 2 to November 13, that this subject was successfully advanced to conditions which seem to exclude all hypotheses of sensory perception. Nevertheless, she still got scores which indicate that something more than chance was involved in the matching.[3]

The procedure in question differed from screened touch matching

[2] Compare with the decline of Stuart, Linzmayer, and others in the Duke experiments.

[3] Various other methods offering more rigid controls than screened touch matching had been tried frequently throughout the experiments, but without success. For example, sealing the cards in envelopes or having the subject "read down through the deck" gave only chance results.

only in the selection, handling, and arrangement of the key-symbol cards. In the earlier procedure the key cards were face up. In this, they were placed face down, special precautions being taken that neither the experimenter nor the subject should sense their order.

As the first step in this direction, the experimenter prepared 28 sets of key-symbol cards in advance. Each set of 5 symbols was shuffled, a blank card was used to cover the face of the bottom symbol to prevent its being accidentally seen at any time, and the 6 cards were fastened together by an elastic band. These 28 sets of key cards were used in rotation, one set being re-shuffled and placed back with the others after being turned face up for checking the results of a series of trials.

A set of receptacles for the place cards was made by fastening five open shallow boxes to a strip of cardboard. These boxes were just large enough to hold the key-cards, but were deep enough for 50 cards. Five blank cards were provided to cover the face-down key cards after they had been placed in the boxes.

In arranging the key cards, the following routine procedure was strictly observed. The experimenter first placed the row of boxes on his knees *under* the table. (Due to the fact that the table was covered with a blanket which hung nearly to the floor, the subject could not have seen the boxes in this position if she had tried.) Then the experimenter selected one of the sets of key cards. Holding them face down under the table, he slid the 5 symbol cards one after another into the 5 boxes. Then the blank cards were placed over the key cards while the boxes were still under the table. Having completed this, the experimenter placed the row of key cards under the edge of the screen before the subject, picked up and cut a deck of shuffled cards, and gave the signal for the subject to begin.

Her responses consisted in pointing, as previously described, to one after another of the key cards, indicating to the experimenter how to place the cards of the deck to match those before her. In this instance, however, the subject was attempting to match cards which she did not see with key cards which she had no sensory way of knowing. It was "blind" matching.

Under the conditions of these tests, the question of shuffling is not so important. That is, the order of the key symbols was unknown, and chance would be expected to yield through a long series a score of about 20 per cent right. As a matter of fact, in these experiments 1,000 shuffled cards were kept at hand, and the trials were rotated

through these by decks of 50 or 100 cards. The experimenter gave each deck 5 dove-tail shuffles immediately after a check-up was made. When a deck was selected for use, it was always cut by the experimenter, out of the subject's sight, just before the signal to start was given.

When the deck was sorted out, the screen was removed, the experimenter turned up the key cards and recorded their order, and the actual distribution of the failures and successes under each key symbol was recorded and checked by both the subject and the experimenter.

This condition was first used on October 30. Between that time and November 9, the subject made 7,800 trials under this condition; 22.4 per cent of her trials were correct. The positive deviation of her hits from chance is 188, while the standard error for this number of trials is 35.3. This gives a critical ratio of 5.3, with odds of 2,500,000 to 1 that the result is not one of mere chance coincidence.

At the time of making this experiment, the subject's scoring ability under all conditions was on the wane. In the hope of restoring her confidence, therefore, the experimenter permitted her to work entirely at the original screened touch matching procedure, key cards face up, for a period of time beginning November 10. However, her decline was not checked, and we have seen that the fourth period yielded results approximating mean chance expectation. Later on during this fourth period, the face-down (or blind) procedure was tried again but this time, as with the other procedure, it yielded only chance averages. This condition was used for 13,700 trials of the 28,775 made from November 14 to December 19 with scores at the 20 per cent level.

In coming back to the blind matching procedure after failure with the regular screened touch matching, there was no reason to expect a return to good scoring. This condition is in fact considered by most as harder, or at least more advanced. Woodruff and George report in this issue that they obtained poorer results with the key cards face-down (blind matching) than with them face-up.

The aim in resuming the advanced condition was a mixed one: in part to check the adequacy of the conditions by a "control" series made when the subject was evidently not capable of ESP. Also, there was the remote possibility that variety might stimulate or challenge new interest and arouse the subject to good scoring again. Finally, it seemed advisable to continue for some time to observe any trend that might appear in the results, negative or positive. As a matter of fact the scores adhered close to the expected mean, evidencing only the play of chance factors.

It would violate the purpose of the experiment to lump together the first and second series of the blind matching work (with key cards face-down). But even if this be done a critical ratio of 3.2 is obtained, still giving odds of 3,000 to one against the chance hypothesis.

As a further step to insure the work against the alternative of chance, 7,500 trials were made with all essential factors present excepting ESP. The aim was to give a close empirical control on the theoretical probability. The procedure as follows: A series of 15 sets of five key cards were chosen in a routine manner and recorded. Then in the same routine way (to avoid ESP) 20 packs of ESP cards were taken as they came from the stock room and two at a time (50 cards) laid down opposite one set of the key cards. The cards were dealt in a routine manner from left to right. Since the matching was uneven in the regular tests, there was a systematic variation introduced in the control to give unequal numbers opposite the key cards. This variation occurred in an orderly way, however, and the whole procedure was mechanized. Only ESP was ruled out.

The percentage of hits in this empirical control series was 19.7 or only .3 of one per cent off the theoretical mean. If we compare now the percentage of 22.4 obtained in the 7,800 trials where ESP was expected to operate, we find a difference that is 4.22 times the standard deviation of the difference, and this is obviously a very significant figure.

If we may, as the results indicate to be necessary, conclude that Mrs. M. at one time demonstrated ESP ability, the decline of the capacity to demonstrate it remains a puzzling problem. But this is in line with the work of others and bids fair to become a general law for routine laboratory tests. Whether the loss will be in this instance permanent remains for further experimentation to discover.

Summary: A program of research has been briefly described which has as its main objectives: (1) further evidence of the occurrence of ESP; (2) further understanding of its nature. The present paper has presented the results of the program only as they bear at present upon the first objective. One subject has been discovered who has given extra-chance scores through many thousands of trials, showing, however, a gradual decline to a chance level of scoring.

One condition of testing which was designed to eliminate all hypotheses of extra-chance scoring except that of ESP was described. These data were selected solely on the basis of the rigor of the conditions of testing; the actual percentage of success, for the selected data

was *lower* than the subject's general level of scoring for all conditions. Considering all the trials in one unbroken series made under this condition, the odds that the result obtained was not mere chance coincidence were 2,500,000 to 1. The conclusion indicated is that the subject was demonstrating in her matching a degree of knowledge of the cards which she could not have acquired through known sensory channels.

EXPERIMENTS IN EXTRA-SENSORY PERCEPTION

J. L. WOODRUFF and R. W. GEORGE

ABSTRACT: Seventeen persons were tested for ability in extra-sensory perception. Three of these were selected for a comparison of success with the following methods: (1) Open matching, in which subject attempts to match cards face down to the appropriate symbol visible before him; (2) blind matching, in which he tries to match the cards to invisible symbols on cards face down before the subject; and (3) verbal calling, in which he names each card of the pack face down on the table. These methods were used also with the cards totally screened from the subject's view, making six conditions in all. The chief aim was to discover whether a test involving conscious attention to the symbols was more successful than a more motor test or was less so. In all the methods high enough scores were made to indicate that extra-sensory perception was operative. The unscreened work produced, with two subjects, better results than the screened work. But the opposite was true for the third subject. The open matching method yielded the highest total score, the screened blind matching the lowest total score.

A block of 3,000 trials was studied to note whether a method excluding extra-sensory perception would yield scores as good as those made by the subjects, and a cross-check of 11,875 trials was made as a further control on the mathematical treatment. The results of both empirical controls were, however, very close to what might be expected by chance alone.

The research was conducted at Tarkio College, where Mr. Woodruff was a senior student in psychology and Dr. George is Professor of Psychology.

THE FOLLOWING REPORT of experiments in extra-sensory perception involves two series.[1] The first series consisted in applying standard techniques to discover, if possible, extra-sensory ability in a number of subjects. The second series was concerned with the determination of possible relations between methods used in testing for this capacity and the scoring rates of selected subjects. As the methods used involve the same material but different perceptual tasks for the subject, the effect of these differences should be reflected in the scoring rate.

[1] The work was suggested by Dr. J. B. Rhine's *Extra-Sensory Perception*, Bruce Humphries Inc., Boston, 1935. This research was undertaken as a project in an Experimental Psychology course at Tarkio College, Tarkio, Missouri, and was continued over a period from January 1935 to June 1936.



Screened Touch Matching, the Technique Used by J. G. Pratt.



Screened Open Matching, One of the Techniques Used by Woodruff and George.

UNIV. OF MICH.

More specifically, the purpose was to find out whether a method involving a more motor judgment and response would favor or disfavor high scoring in comparison with a method involving a more conscious cognitive one. This problem arises in a general way from the frequent association of claims of extra-sensory perception with automatic movements, but it is more specifically suggested by the work of Tyrrell, in which a subject is found to demonstrate ability in extra-sensory perception using a method relying upon motor responses. The motor responses in Tyrrell's experiments consisted in opening the chosen one of five small boxes, the aim being to choose that in which a lamp was lit. Tyrrell urged the superiority of this motor response over the more cognitive response such as choice of one of five symbols would entail. An experimental testing of Tyrrell's hypothesis was desirable.[2]

METHODS

The regular ESP cards were used for the entire period of experimentation. The packs contain twenty-five cards with five cards of each of five designs.

In this case, then, a score of five correct hits out of twenty-five cards is the chance average, and anything above that over a sufficient number of trials may be regarded as indicative that some factor other than chance has entered into the production of the scores.

In this work there were three distinct methods of procedure employed, and each method was further varied by the use of a screen in approximately half the tests. There were, therefore, six different sets of scores for each subject.

The following symbols will denote the various techniques used: OM, SOM, BM, SBM, BT, SBT. OM denotes the ordinary matching work; i.e., five target cards, one each of the five symbols, are placed face up on the table, and the subject attempts to place opposite the appropriate target cards, one at a time, the twenty-five cards of a pack held in his hand. The cards laid down in the matching remain in a face down position until the check is made at the end of the run. SOM denotes OM work with the packs of cards to be placed screened from the subject's sight by means of a horizontal wooden screen. BM is that work done with both the target cards and the deck of cards face down. In this technique no attempt is made by the subject consciously to ascertain the symbols on either the target cards or the deck

[2] G. N. M. Tyrrell, "Further Experiments in Extra-Sensory Perception," *Proceedings of the Society for Psychical Research*, XLIV, 99-167. See article on Tyrrell's work in this number by Burke M. Smith.

of cards; his only task being to judge to which target a card taken from a pack is similar. BT indicates that the subject names aloud each of the twenty-five cards as he lifts it from the deck. An observer records each call. After the complete deck of twenty-five has been run through in this manner, the calls made by the subject are checked with the actual order of the cards. The letter S used in connection with the above symbols indicates that the conditions include a screen which blocks the subject's sight of the cards he handles.

The general experimental conditions are of particular importance in the conduct and evaluation of work of this kind. The greatest care was taken to assure their reliability. New packs of cards were used frequently, and those showing any signs of soil or wear were discarded. With the exception of 2,250 trials run by the experimenter J. W. himself, unobserved, all the trials were witnessed. After each run of twenty-five trials, and in no case before twenty-five had been run, the scores were checked by the observer. After each series the observer shuffled the cards; the same deck of cards was not used for any two runs in succession, but three or four decks were used in rotation. The screen used in the screened series was a low four-legged table, approximately six inches high and two feet square. This was placed on another higher, larger table. In the SOM and SBM techniques, the cards were shuffled by the observer, then passed under the screen to the subject. The target cards were placed on top of the screen and the subject then matched them by placing his deck, card by card on the lower table, under what he thought was the appropriate target card. In the SBT technique the subject named the screened cards as he lifted each one from the deck. The subject's hands and the cards held by him were hidden from his sight by the upper table throughout.

Where it is possible the results have been shown in tables. The notation "Dev. from np" denotes the difference between the number of hits scored and the chance expectation. X (= Dev./S.D.) denotes the critical ratio computed by dividing this difference by the Standard Deviation of the mean expectation. An X value of 2.5 or greater, which indicates odds of about 162 to 1 that some other factor than chance is responsible for the deviation, is customarily considered significant. This criterion is approximately equal to the 4 P.E. used by Rhine.[3]

RESULTS OF FIRST EXPERIMENTAL SERIES

The BT technique was used exclusively during the first period of experimentation (January 1935-June 1935). The subject was at

[3] *Op. cit.*, p. 32.

liberty to look at the backs of the cards if he desired, but many runs were made with the subject's eyes closed. However, no accurate division can be made as to the number of trials run in this way. Table One gives a summary of this early work. It is evident that the significant results are due mainly to the scores made by J. W.

TABLE ONE

WORK COMPLETED BETWEEN JANUARY 1935 AND JUNE 1935. BT METHOD.

Subject	Number of Trials	Number of Hits	Deviation from np	Average per 25.
W (Obs.)	2175	559	124	6.4
W (Unobs.)	1750	527	177	7.5
C	1675	360	25	5.3
M	1000	211	11	5.3
S	500	94	— 6	4.7
E	300	61	1	5.1
G	250	55	5	5.5
T	250	52	2	5.2
D	250	49	— 1	4.9
Totals	8,150	1,968	338	6.0 X = 9.36

RESULTS OF SECOND EXPERIMENTAL SERIES

In the fall of 1935 an exploratory series of tests was made to aid in selection of suitable subjects for experimentation to follow. The method was again BT, and the subject was allowed to continue the tests as long as he desired. However, those who were scoring highest were *urged* to continue with their work. The fact that no subject was limited in the number of trials he might attempt, even though his apparent success was negligible, possibly caused the general average to run lower than it otherwise might have. This preliminary testing was conducted over a period of two months. In Table Two is the summary of this work, all of which was done with the BT procedure. Again the results are well over the most exacting criteria for extra-chance events. This time four of the 16 subjects yielded significant averages, with others close.

RESULTS OF THIRD EXPERIMENTAL SERIES

From the above candidates, J.B., M.S., R.C., and J.W. were selected as subjects for the remainder of the experimentation. R.C. did not enter school the second semester; so the work proceeded from that point with only three subjects.

Over a period from November 1935 through April 1936, technique comparisons were made in three sets, the results of which are given in Table Three. From November to January OM, BM, and their screened

TABLE TWO

BT WORK: WOODRUFF OBSERVER: SEPT. 1, 1935-OCT. 31, 1935. TEST FOR SUBJECTS.

Subject	Number of Trials	Number of Hits	Deviation from np	Average per 25.
F. L.	125	25	0	5.00
M. S.	500	131	31	6.55
R. E.	250	64	14	6.40
E. P.	500	103	3	5.15
R. M.	500	102	2	5.10
W. S.	250	52	2	5.20
J. M.	500	93	— 7	4.65
H. P.	250	50	0	5.00
V. H.	500	119	19	5.95
A. C.	250	55	5	5.50
L. T.	225	38	— 7	4.20
M. W.	50	9	— 1	4.92
R. C.	1,550	351	41	5.66
J. B.	1,500	351	51	5.85
W. G.	250	36	—14	3.60
J. W. (Unobs.)	500	135	35	6.75
J. W.	2,800	704	144	6.28
Totals	10,500	2,148	318	5.75 X = 7.51

TABLE THREE

COMPARISON OF OM TECHNIQUES WITH BM TECHNIQUES.

Name of Subject	OM			SOM			BM			SBM		
	Tr.	Sc.	Av.	Tr.	Sc.	Av.	Tr.	Sc.	Av.	Tr.	Sc.	Av.
D. C.*	250	51	5.10	0	0	.00	75	8	2.80	0	0	.00
R. C.	600	118	4.91	350	70	5.00	350	63	4.50	250	42	4.20
J. B.	1500	386	6.43	1050	225	5.35	1300	272	5.23	800	185	5.78
M. S.	1025	277	6.75	975	289	7.41	75	185	5.28	700	164	5.85
J. W.	3450	1124	8.14	3100	702	5.66	3325	739	5.55	2625	563	5.36
Totals	6825	1956	7.16	5475	1286	5.86	5925	1267	5.34	4375	954	5.45

Grand Totals—Trials 22,600; Score 5463; Dev. 943; Av. 6.04; X = 15.6

COMPARISON OF OM TECHNIQUES WITH BT TECHNIQUES.

Name of Subject	OM			SOM			BT			SBT		
	Tr.	Sc.	Av.	Tr.	Sc.	Av.	Tr.	Sc.	Av.	Tr.	Sc.	Av.
R. C.	125	29	5.80	125	21	4.20	125	22	4.40	125	29	5.80
J. B.	1200	241	5.02	1200	299	6.22	1200	269	5.60	1200	294	6.12
M. S.	625	150	6.00	625	227	9.08	625	141	5.64	625	213	8.52
J. W.	2200	683	7.77	2200	487	5.53	2200	703	7.98	2200	472	5.36
Totals	4150	1103	6.65	4150	1034	6.22	4150	1135	6.83	4150	1008	6.07

Grand Totals—Trials 16,600; Score 4,280; Dev. 960; Av. 6.44; X = 18.6

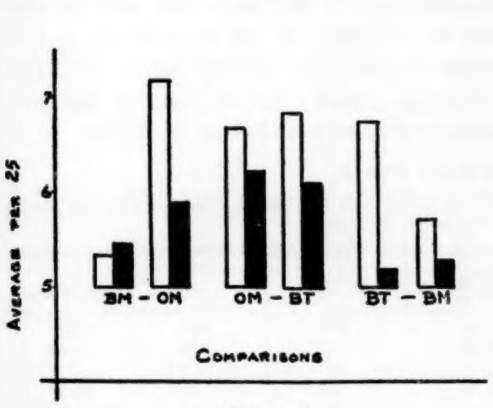
COMPARISON OF BM TECHNIQUES WITH BT TECHNIQUES.

Name of Subject	BM			SBM			BT			SBT		
	Tr.	Sc.	Av.	Tr.	Sc.	Av.	Tr.	Sc.	Av.	Tr.	Sc.	Av.
J. B.	1150	242	5.26	1150	237	5.15	1150	250	5.43	1150	223	4.84
M. S.	225	47	5.22	225	52	5.77	225	41	4.55	125	22	4.40
J. W.	1400	341	6.08	1400	296	5.28	1400	458	8.14	1400	310	5.53
Totals	2775	630	5.68	2775	585	5.27	2775	749	6.74	2675	555	5.18

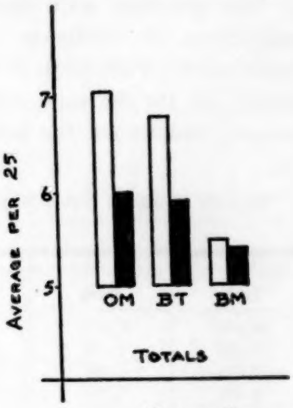
Grand Totals—Trials 11,000; Score 2519; Dev. 319; Av. 5.72; X = 7.16

*Served as subject in comparison of techniques for only 325 trials. Had no preliminary test period.

variations were compared. From January through February, OM, BT, and their screened variations were compared. March and April were devoted to the comparison of BM, BT, and their screened procedures. The results are represented in Graphs 1 and 2.



GRAPH 1



GRAPH 2

GRAPH 1. Results of the paired comparisons of the various techniques. The white column represents the total average above chance per 25 made under the unscreened conditions of the technique; the black column indicates the average when a screen is interposed.

GRAPH 2. The techniques totaled and compared with their screened conditions.

It is evident on examination of Table Three that each group of scores for each technique greatly surpasses the accepted standard of significance (2.5 times D/S.D.). Since the conditions were such that the extra-chance scores are attributable to no sensory mode of dealing with the card symbols, it would follow that these subjects have demonstrated ability in extra-sensory perception.

The problem, however, that primarily concerns us is, as has been stated earlier, the comparison of the OM and BT techniques with the BM technique. In the first two types the subject pays conscious attention to the symbols on the cards, while in BM work he makes no consciously directed effort to determine the symbols; his task is simply to recognize similarity between the symbols on the deck of cards and the target cards. In other words, he matches the deck of cards with the five target cards, placed face down in an unknown order, guided only by a "feeling" that the symbol on one card is similar to the symbol on another.

The OM in comparison with BT offers the control needed on a manual response versus an oral. If a manual response is better, OM and BM should excel over BT. If inattention to the symbols helps, BM alone should stand out. If conscious attention is favorable, OM and BT should range highest.

The subjects, with the exception of J.W., were not told of the comparison of technique scores to be made at the completion of the experiment. Following is a table showing the relative scores. All the results are for the same experimental period; that is, none of the preliminary results for the testing periods are included in the tables.

TABLE FOUR

SCORES MADE BY FOUR SUBJECTS FOR ENTIRE PERIOD OF WORK EXCLUSIVE OF TESTING PERIOD.

Trials	Score	Average per 25.	Deviation from np	X	Type of Work
10,975	3,073	7.08	878	20.9	OM
9,625	2,320	6.02	395	10.0	SOM
6,925	1,884	6.79	499	14.9	BT
6,825	1,563	5.72	198	5.9	SBT
8,700	1,897	5.45	147	3.9	BM *
7,150	1,539	5.38	109	3.2	SBM *
50,200	12,276	6.11	2,226	24.95	All types

*No conscious attention.

The above results seem to point rather strongly to the importance of conscious attention to specific symbols in ESP. In every case both the averages and the X values are higher for the conscious attention technique (OM and BT) than in BM work where conscious attention to the symbols is not so noticeably operative. However, it must be recognized that even in the BM and SBM techniques the X value is above the criterion for significance (2.5).

It is necessary, however, to recognize an alternative hypothesis for the explanation of these results—an hypothesis which seems to have

some possibilities as an explanation. Without exception the subjects stated that the perceptual judgment task involved in BM work was more difficult than that involved in the other techniques. The results, then, may possibly be explained by the fact that low scores seem to result more frequently when the task *appears* difficult.

EFFECT OF SCREENING

In general, as Graph 1 and Table Three will show, the screens were inhibitory, but with the noteworthy exception of subject M.S. He improved with the screen and his reports indicate that the unscreened deck of cards distracted him. On the other hand, J.W. felt that the presence of the cards helped him to concentrate attention. For these individual comparisons see Table Five, and for general screen effects, Graphs 1 and 2.

TABLE FIVE

COMPARISON OF INDIVIDUAL SCORES WITH RESPECT TO TECHNIQUES AND THEIR SCREENED CONDITIONS.

SUBJECT: R. C.

OM		SOM		BM		SBM		BT		SBT	
Tr.	Av.	Tr.	Av.	Tr.	Av.	Tr.	Av.	Tr.	Av.	Tr.	Av.
725	5.07	475	4.78	350	4.64	250	4.20	125	5.80	125	4.40

Trials 2,050; Score 394; Average 4.80; Dev. —16; X = 0.88

SUBJECT: J. B.

OM		SOM		BM		SBM		BT		SBT	
Tr.	Av.	Tr.	Av.	Tr.	Av.	Tr.	Av.	Tr.	Av.	Tr.	Av.
2700	5.93	2250	5.82	2450	5.24	1950	5.41	2350	5.52	2350	5.39

Trials 14,050; Score 3,137; Average 5.56; Dev. 317; X = 6.89

SUBJECT: M. S.

OM		SOM		BM		SBM		BT		SBT	
Tr.	Av.	Tr.	Av.	Tr.	Av.	Tr.	Av.	Tr.	Av.	Tr.	Av.
1650	6.46	1600	8.06	1100	5.25	925	5.84	850	5.35	750	7.83

Trials 6,875; Score 1,808; Average 6.56; Dev. 433; X = 13.05

SUBJECT: J. W.

OM		SOM		BM		SBM		BT		SBT	
Tr.	Av.	Tr.	Av.	Tr.	Av.	Tr.	Av.	Tr.	Av.	Tr.	Av.
5650	7.99	5300	5.60	4725	5,71	4075	5.33	3600	8.06	3600	5.43

Trials 26,900; Score 6,878; Average 6.38; Dev. 1,498; X = 22.83

It is interesting to note that in BT that technique which most depends upon conscious awareness of the symbols, which is perhaps most cognitive in type, the screening of the cards had the greatest inhibition of scoring. Next, in OM, where one need not use the imagination to recall the symbols—they are before his eyes—screening is next in its inhibitory effect. Finally in BM, where the key cards and the trial cards are already face down and thus in effect already screened, the screen had its least effect. The decrease in scoring with the introduction of the screen was BT, 60 per cent; OM, 51 per cent; BM, 16 per cent.

TABLE SIX

SCORES DURING FEBRUARY 1936 PERIOD WHEN REWARDS WERE OFFERED.

Subject	OM			SOM			BT			SBT		
	Tr.	Sc.	Av.	Tr.	Sc.	Av.	Tr.	Sc.	Av.	Tr.	Sc.	Av.
M. S.	375	97	6.46	375	130	8.66	375	89	5.93	3.75	117	7.80
J. B.	875	180	5.14	875	226	6.45	875	184	5.25	875	217	6.20
J. W.	1250	435	8.70	1250	276	5.52	1250	425	8.5	1250	274	5.40
Totals	2500	712	7.12	2500	632	6.32	2500	698	6.98	2500	608	6.08

Grand Total—Trials 10,000; Score 2,650; Av. 6.6; Dev. 650; X = 16.25

It may be noticed in Table Six that in the month of February the X values and averages are higher than during other periods. During this period an effort was being made to ascertain, if possible, what effect a special reward might have on the scoring efforts of the subjects. On alternate weeks during the month, J.B. and M.S. were given the opportunity of competing for a small prize (a theatre ticket) which went to the subject making the highest average score during the week. A comparison of the scoring levels on the alternate weeks brings little of significance to light. In the OM procedure the scores for the reward weeks were lower. In the SOM technique M.S.'s scores declined throughout the month regardless of the reward. In the BT work there was no noticeable trend. In SBT a steady improvement occurred throughout the month. These results would seem to indicate that the reward had no effect upon scoring, but when February's results are compared with those of other months an interesting difference appears. In January all trials gave an X value of 9.54; this value rose to 16.25 in February when the reward was offered and fell to 6.3 in March when the rewards were discontinued. Approximately the same number of trials were made each month.

Table Seven shows the averages of the subjects on the alternate weeks of the month. The first and third weeks in the month are those

in which a reward was offered. Approximately the same number of trials were run by each subject with the exception of the fourth week when M.S. ran no trials.

TABLE SEVEN

SHOWING AVERAGES PER 25 BY WEEKS THROUGH FEBRUARY, COMPARING REWARD WITH ABSENCE OF REWARD.

Subject	OM	SOM	BT	SBT	Remarks
J. B.	5.2	5.5	5.3	5.2	Reward
M. S.	5.2	9.0	6.8	6.6	
J. B.	6.1	6.0	5.5	5.4	No reward
M. S.	9.4	8.8	5.6	7.4	
J. B.	4.1	8.9	5.4	6.8	Reward
M. S.	4.8	7.8	5.4	9.4	
J. B.	5.1	5.8	4.6	6.7	No reward
M. S.	No trials				

Although the reward had no very specific effect indicated in the results, there remains the possibility that a general increase in interest owing to the reward situation may have accounted for the total higher scoring during the period.

The results of all work completed during the entire period of experimentation are summarized in Table Eight. The odds against this deviation being attributable to chance are 10^{166} to one.

TABLE EIGHT

COMPILATION OF ALL WORK OF ALL SUBJECTS JANUARY 1935-JUNE 1936.

Trials 68,850; Score 16,662; Average 6.05; Dev. from np 2,892; X = 27.55.

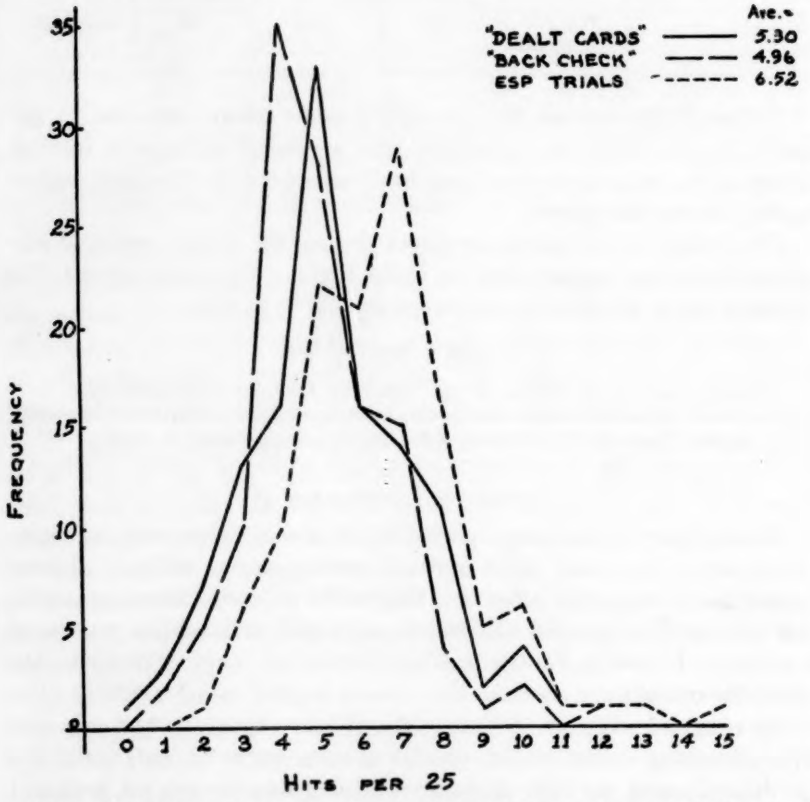
SUBJECT'S ESTIMATION

During part of the second period an incidental experiment was conducted which was based upon introspective reports of subjects who indicated that it was their belief that they could estimate rather accurately their scores. For a period of a month and a half each subject was asked to estimate, following the completion of a run of twenty-five cards, but before the cards were checked, the number he had called correctly. The results obtained were so obviously barren when graphed that only one correlation was made, which yielded a correlation of only .025 for J.W.'s series with the OM method. The fact that the subject hesitated to estimate a very high or a very low score probably reacted against a high correlation. The estimating of scores was discontinued because the results were so fruitless, and because it tended to make the sub-

jects "score conscious," an effect which seems to lower scoring considerably.

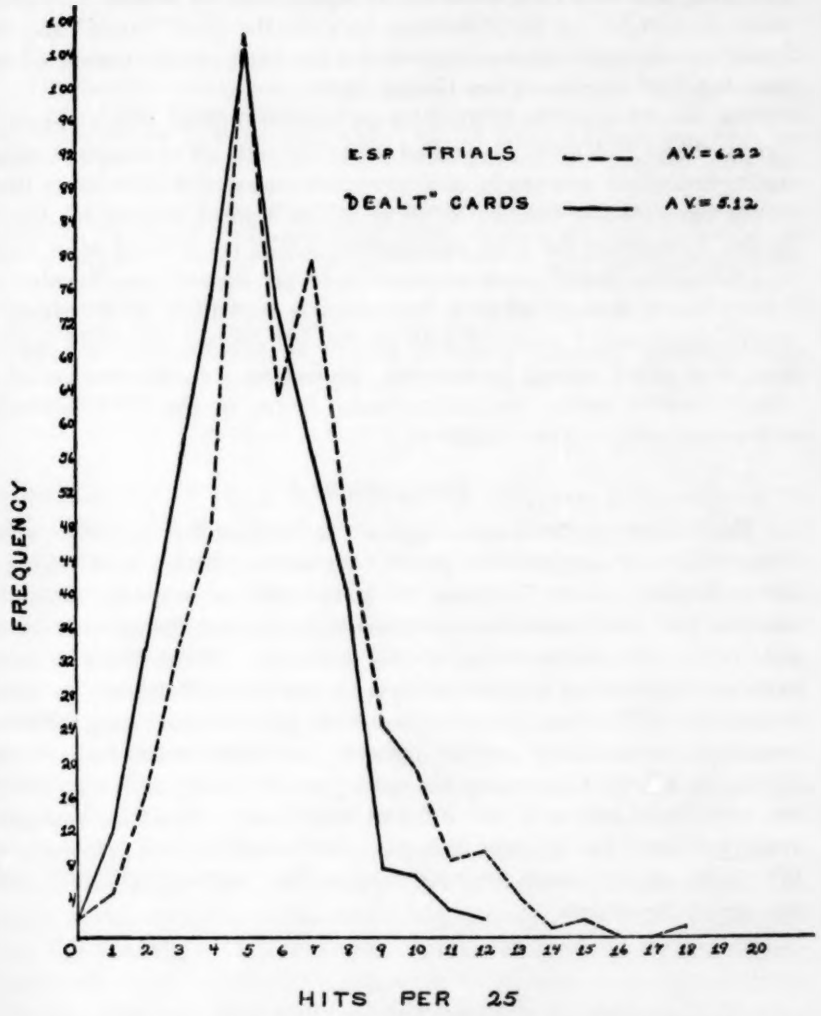
EMPIRICAL CHECKS

During the year, two other projects were carried out which were occasioned largely because of suggestion made by some critics to the effect that if control series were run parallel with experimental series, the resultant curves would in all likelihood be very similar. In the first of these a curve was drawn from the last 3,000 of the BT trials completed. As a control series the calls made by the subject by runs of 25 during the 3,000 trials were checked against the order of the cards in the preceding run of 25 cards. In this way it was possible to make an



GRAPH 3. A comparison of the hit-frequencies obtained when a shuffled pack of cards was matched to a fixed pattern, when the call series were checked against another than the intended card series, and when the call series were checked against the card series for which they were intended.

empirical check within the same material. The test series and the control series differ only in that in the test series the subject's calls are matched against the cards he desires them to match, and in the control series they are matched against similar series of cards which have no relation to his striving. In order that a further empirical check might be made, 3,000 cards were dealt in rotation, the dealer placing them



GRAPH 4. A comparison of the hit-frequencies observed in 11,875 ESP trials, and the hit-frequencies resulting when the same number of cards were dealt out in a fixed manner.

under the target cards, the order of which was not changed. The number of hits was checked after each series of 25. The average for ESP trials is 6.52 hits per 25. The average for the "backcheck" series was 4.96 per 25. The average per 25 for the "dealt" series was 5.3. The difference in averages per 25 between the ESP series and the "backcheck" series is $1.56 \pm .2583$ which gives an X value of 6.04, indicating this difference to be highly significant. A similar significant value, $X = 4.72$, for the difference between the ESP "trials" and the "dealt" cards, again shows emphatically the extra-chance nature of the intended ESP series. (See Graph 3).

The second experiment made as an empirical check was a comparison of the 11,875 trials (including all trials with all techniques) up to the time the test was made, and a control series of 11,875 cards dealt in the same routine manner as those in the control experiment above. In this comparison the ESP trials averaged 6.22 per 25 with an X value of 13.3. The "dealt" cards averaged 5.12 per 25 with an X value of 1.35. The difference between the averages is $1.10 \pm .1298$, giving a highly significant X value of 8.47 for the difference. It would appear that, even when chance factors are doubled by the introduction of a chance control series, the extra-chance factor in the ESP series is still outstanding. (See Graph 4).

CONCLUSIONS

The following conclusions appear to be justified by this work. First, it further confirms the already extensive evidence that clairvoyant perception occurs. Second, it seems safe to say that with the subjects here used, conscious attention is important, though not essential, in the efficient operation of this principle. Third, there is some evidence that at least certain subjects do not respond directly to small rewards in ESP tests. Fourth, the ability to perform under different conditions varies widely and in different directions among individuals. Fifth, the effect of screening the cards varied widely with the different individuals and with the different techniques. Its effect was generally inhibitory but this was least with blind matching and greatest with BT. One subject clearly improved and another markedly declined with the use of the screen.

AN EXPERIMENT IN CARD GUESSING

C. R. CARPENTER and HAROLD R. PHALEN

ABSTRACT: Twenty-four college students were tested with ESP form cards. Two were found whose averages were significantly above chance. Cards bearing colors instead of forms were used. Out of thirty subjects, four were successful in scoring significantly. These four included the two previously successful subjects. The colors were generally as effective as the symbols. Three of the successful subjects were carried forward to tests involving greater precautions. Subject G made significant scores on three of four further experiments. Subject L made significant scores on two of three further experiments. Subject M failed to make significant scores after the techniques were changed. The selected subjects' results were compared with the scores of ten unsuccessful subjects, and with a check series of 10,000 trials.

The research was conducted at Bard College of Columbia University where Dr. Carpenter is Fellow and Lecturer in Psychology and Dr. Phalen is Professor of Mathematics.

THE FINDINGS of Dr. J. B. Rhine and his associates as reported in his now well known monograph need further verification, explanation, and interpretation. By means of objective procedures he has obtained results which are inconsistent with currently accepted views. It would seem to be obligatory for scientists to consider these data carefully and to criticize them strictly either favorably or unfavorably. The findings are not readily harmonized with established principles in subjects of perception and learning, stimulus and response, and the recognized processes of communication. It is significant that the conclusions of the workers in parapsychology are being accepted by influential scientists other than psychologists.

Reports, published and unpublished, indicate that the numerical findings of Rhine and his workers are subject to corroboration; this paper furnishes further evidence on this question. By following in general the procedures and statistical techniques outlined by Dr. Rhine, numerical results which are comparable to his have been secured after a large number of trials with a few of the subjects tested.

We felt the challenge offered by Rhine's findings and interpretations just as the challenge has been felt by other psychologists. With-

out identifying ourselves with either the supporters or the opponents of the work in parapsychology we have considered it permissible to conduct carefully planned experiments relative to this significant problem in psychophysics and to report our findings as impartially, definitely and clearly as our statistical procedures permit. In the very beginning, let us state clearly that we do not intend to draw conclusions relative to the validity of citing "telepathy" and "clairvoyance" as *explanations* of the observed findings, that cards are guessed with a score that is better than chance expectancy. In this respect we are inclined to await further experimental analysis and to consider other possible interpretations. *In this paper we shall simply present our numerical findings for the consideration and analysis of the interested reader.*

It was with an attitude of skepticism regarding Rhine's conclusions that we began in 1934 a series of experiments using his procedures as nearly as possible in order that we might check some of the factors believed to be operating to produce his results. It was thought at that time that errors might be found either in his experimental procedures, in his statistical methods, or in the validity of his generalizations. There was serious doubt as to whether or not we could find individuals with the ESP (extra-sensory perception) ability when using thoroughly objective and strictly controlled experimental techniques.

The specific purposes of the work undertaken were: 1. To set up experiments similar to those of Rhine and to ascertain whether or not his findings could be duplicated in another laboratory with other workers and subjects. 2. To learn whether or not other kinds of stimuli, e. g., color cards, than the ESP cards could be used with equal or greater accuracy by individuals who demonstrated the ESP ability on these cards. 3. To learn the percentage of a relatively large number of individuals who could guess cards with an accuracy above chance expectancy when the subjects were tested in objective, uniform situations.

Subjects. The subjects used in these experiments were men who ranged in age from seventeen to twenty-six years. In addition two young women were used. In all a total of 100,000 trials were made (a trial equals one attempt to guess a card correctly) during eight different series of experiments using slightly different methods. With the exception of the two women who were faculty wives, the subjects were students of Bard College of Columbia University and represented the four college classes. They came from all branches of study, including the physical, biological and social sciences, mathematics, the languages, literature, art, music and drama.

General Procedures. In order to eliminate the selective factor as much as possible, subjects were secured by student assistants without the use of persuasion. Subjects were not paid for their time or informed that they would be paid during the exploratory series of trials. However, three of the "gifted" subjects were paid at the rate of thirty-five cents per hour during intensive periods in order that we might have enough of their time for effective work. When subjects came for the experimental work, they were told briefly about the procedures and the object of the experiment. They were asked to familiarize themselves with the stimulus cards, either the ESP cards with forms or the cards with colors. The deck contained in all cases either five cards each of the ESP forms (having symbols, O □ + ≋ ★, as used by Rhine) or five cards each of the colors, red, yellow, green, white and blue. The subjects were told that the distribution would be random, since the deck would be shuffled thoroughly, and that they were to guess each card so as to make a score that was better than that to be expected by chance, i. e., on the average five correct guesses per twenty-five trials. The only efforts to establish rapport with the subjects were those common to any experiment in which it is necessary to have the co-operation of the subject.

The N(Normal) Series. It was considered desirable to have a completely objective series of trials from which the human factors were eliminated. Rather than rely solely upon the theoretical possibilities of curve construction, we wished to construct a comparative curve on the basis of our empirical data. Therefore, we ran a series of 10,000 trials involving only chance. Two decks of ESP cards were used. Both decks were cut and shuffled at least five times. The distribution of the target deck was recorded against the distribution of the matching deck and the number of hits or corresponding cards was checked and summed for each series of twenty-five trials.

Results of the N Series. The average number of hits for each series of twenty-five trials was found to be 5.03. The mode fell at 4. The number of correct hits per twenty-five trials ranged from 1 to 11, there being 11 instances of 1 hit per twenty-five trials and 1 instance of 11 hits. It is to be noted that the distribution is not perfectly symmetrical (Graph I) since there is a skewness to the left toward the lower numbers.

The result of the matching of the two packs of cards is a curve of coincidence approximating the theoretical distribution. In both distributions of both packs it is certain that there will be an equal num-

ber of suits and of individual cards in the suits. The distributions are random and the hits occur with random coincidence.

Series 1. The decks of cards previously described were changed frequently, new decks being employed when used decks became soiled, bent or otherwise marked so that there was a possibility of the subject's getting cues, even though unconsciously, from the backs of the cards. The deck to be used in a particular series of twenty-five trials was shuffled at least five times and cut. It was handed to the subject who held it in his hands. He called the bottom card of the deck, removed it, placed the card face down on the table without observing whether or not he had called correctly. He then guessed the symbol of the next card, removed that and placed the card face down on top of the first, and thus he continued until the entire pack of twenty-five cards had been guessed. (Calling the bottom card has the advantage of preventing the subject from seeing the card but on the other hand, if the work is done over a smooth surface, there is the possibility of reflections giving cues as to the form stimuli.)

As the guesses were made, they were recorded by the experimenter. When the deck was finished, the subject turned it over, face up and read off the true distribution to the experimenter. The calls or guesses, therefore, were recorded in a parallel column alongside corresponding cards of the true distribution. The number of hits or correct guesses was checked by the experimenter and the subject informed of the score on the twenty-five trials. Also, at the end of each two hundred and fifty trials, the score was averaged and this was reported to the subject. This periodic reporting of scores kept the subject interested and well motivated for the rather monotonous task of guessing and reading cards. The pack of cards used in one series of twenty-five trials was set aside and another deck used for the following series and thus the decks were alternated throughout the experiments.

In all experiments marked Series 1, this procedure was held constant with the following exceptions: Periodically the experimenter and the subject were observed by a third person who checked the controls and scoring. This second observer was especially concerned to see that the subject handled his cards in such a manner as to prevent his getting visual cues directly from the stimuli. Obviously tactual cues were not controlled in this procedure.

Results of Series 1. For these data the value of the probable error was obtained from the formula

$$\text{p. e.} = .6745 \sqrt{npq}$$

TABLE 1.

RESULTS OF 29,750 TRIALS ON SERIES 2—1935.

Number	Name	Number Trials	Number Hits	Average Per 25.	Dev.	P. E.	X
1.....	Roc.	3000	632	5.29	32	14.78	2.16
2.....	Pat.	2000	417	5.21	17	12.07	1.41
3.....	Kri.	1750	338	4.83	—12	11.29	1.04
4.....	Wil.	1000	195	4.88	— 5	8.53	.58
5.....	Koe.	1000	206	5.15	6	8.53	.70
6.....	Ter.	1000	190	4.75	—10	8.53	1.17
7.....	Mur.	1000	188	4.70	—12	8.53	1.40
8.....	Leo.	1000	193	4.83	— 7	8.53	.82
9.....	Moy.	1000	189	4.73	—11	8.53	1.29
10.....	Bai.	1000	201	5.03	1	8.53	.11
11.....	Fic.	1000	200	4.75	0	8.53	0
12.....	Eas.	1000	199	4.98	— 1	8.53	.11
13.....	Hav.	1000	210	5.250	10	8.53	1.17
14.....	Cor.	1000	203	5.08	3	8.53	.35
15.....	Smi.	1000	213	5.33	13	8.53	1.54
16.....	Zel.	1000	212	5.30	12	8.53	1.40
17.....	LaB.	1000	189	4.73	—11	8.53	1.29
18.....	Jon.	1000	201	5.03	1	8.53	.11
19.....	Sim.	1000	182	4.63	—18	8.53	2.11
20.....	McM.	1000	222	5.55	22	8.53	2.58
21.....	Fro.	1000	214	5.35	14	8.53	1.64
22.....	Bol.	1000	230	5.75	30	8.53	3.52
23.....	G.	1500	448	7.47	148	10.45	14.16
24.....	M.	2500	643	6.43	143	13.49	10.6
Totals.....	.....	29750	6315	5.31	365	45.53	8.02

where n equals the number of trials, p equals the probability of success on a single draw (i. e., 1/5), q = 4/5. This formula is common in statistical work and was the one used by Rhine. (See Gavett, *First Course in Statistical Method*, pages 150 and 180). From it can readily be derived the relation,

$$\log \text{p. e.} = (9.43104 - 10) - 1/2 \log n.$$

Table 1 gives the results for twenty-four subjects employed in 1935. Using the criterion that a significant score is indicated if D/p. e. is in excess of four, two subjects were found to have made significant scores, namely G and M. One other subject, Bol., made a score worthy of notice but it fell short of the required criterion.

G was given 1,500 trials, made 448 hits, and his critical ratio was found to be 14.16. M had 2,500 trials, called correctly 643, and his critical ratio was calculated to be 10.6. Bol. obtained 3.52 for his value of X. Since a value of 6 for X indicates that the odds are 20,000 to 1 that the result is not pure chance it is evident that the ratios for G and M are most certainly due to some factor other than chance.

Graph I gives the distribution of hits per twenty-five trials on 10,000 trials selected from Series 2. This summary was made to show the curve of correct guesses for subjects who did not make significant scores. It is to be observed that the number of hits per twenty-five trials ranged from 1 to 12, and that the modal point of the distribution fell at 5. The curve corresponds rather closely with a comparable one constructed on the basis of data from Series N, but it was found that the mode falls at five instead of four, that there is a slight skewness, and that there were two hits of twelve per twenty-five which were out of range of the N Series results.

Series 2. The texture of the cards used in this series was similar to that of Series 1. The stimuli on the face of the cards, however, were *colors rather than forms.* The following colors were printed by a screening process on the face of regular bridge cards: red, yellow, green, white and blue. Between the stimulus color and the edges of the card there was a narrow margin of about one-sixteenth of an inch; this margin was white. The edges of the cards were gilt, the backs contained a pattern of print and under these patterns the cards were colored either red or blue. Decks with red and blue backs were used in about equal proportions. There was no possibility of reading through the cards, either in this series or in the experiments with forms as stimuli.

The procedure of guessing by the subject was similar to that of Series 1, except that after being shuffled and cut, the cards were stacked in an even pile and the subject called the *top* one of the deck *before touching it.* Then the card was removed by lifting it by its edges and placing it face down on the table. Presumably the subject had no information as to whether or not he had called correctly or as to the true color. *The subject always made a call or guess of the card before touching it.*

This procedure was designed to guard against two possible errors which were discovered by preliminary trials with one experimenter as subject: 1. If the card was lifted before it was called, there was a slight reflection of light on the surface underneath the card which might give a cue as to its color. 2. If the cards' faces were touched before the subject made his guess, there was a possibility of a cue because of a difference in the texture of the different colored inks. There remained the very limited possibility that some subjects might have been identifying the color of the card after it was called and thus inferring something of the true distribution.

TABLE 2.

SHOWING RESULTS ON 33,350 TRIALS ON SERIES 3—1935 AND 1936.

Number	Year	Subject	Number Trials	Number Hits	Average Per 25.	Dev.	P. E.	X
1.....	1935	Kap.	1000	209	5.23	9	8.53	1.17
2.....	1935	Hav.	1000	204	5.10	4	9.53	.47
3.....	1935	Zel.	1000	216	5.40	16	8.53	1.88
4.....	1935	Spi.	1000	197	4.90	— 3	8.53	.35
5.....	1935	Mar.	1000	195	4.88	— 5	8.53	.86
6.....	1935	Mea.	1000	203	5.08	3	8.53	.35
7.....	1935	Cre.	1000	185	4.62	—15	8.53	1.76
8.....	1935	M.	3250	883	6.80	233	15.38	15.20
9.....	1935	G.	2500	789	7.89	289	13.49	21.40
10.....	1936	Bol.	1000	175	4.38	—25	8.53	2.94
11.....	1936	Erl.	500	98	4.90	— 2	6.03	.33
12.....	1936	Wag.	500	97	4.85	— 3	6.03	.49
13.....	1936	Par.	500	109	5.45	9	6.03	1.49
14.....	1936	Rai.	500	97	4.85	— 3	6.03	.49
15.....	1936	Gre.	500	93	4.65	— 7	6.03	1.16
16.....	1936	Hop.	500	101	5.05	1	6.03	.16
17.....	1936	Kir.	500	103	5.15	3	6.03	.49
18.....	1936	Sut.	500	104	5.20	4	6.03	.66
19.....	1936	Yea.	500	107	5.35	7	6.03	1.16
20.....	1936	Sch.	500	91	4.55	— 9	6.03	1.49
21.....	1936	Whi.	500	96	4.80	— 4	6.03	.66
22.....	1936	Mer.	500	105	5.25	5	6.03	.83
23.....	1936	Sny.	1000	211	5.27	11	8.53	1.29
24.....	1936	Chu.	2350	508	5.41	38	13.08	2.90
25.....	1936	Bur.	500	135	6.75	35	6.03	5.80
26.....	1936	Hon.	500	112	5.60	12	6.03	1.99
27.....	1936	Kno.	500	104	5.45	4	6.03	.66
28.....	1936	San.	500	116	5.80	16	6.03	2.66
29.....	1936	Jac.	500	94	4.70	— 6	6.03	.99
30.....	1936	L.	7750	2116	7.13	566	23.75	23.80
Totals.....	....	....	33350	7853	5.89	1183	49.27	24.01

Results of Series 2. Thirty subjects were given 33,350 trials in this series. M was given 3,250 trials during which he called correctly 883 making an average score of 6.80 and making an X of 15.2. G was given 2,500 trials during which he called correctly 789 cards thus making an average of 7.89 and a critical ratio of 21.4. In 1936 L was carried forward to this series and given 3,000 trials. He scored 671 hits and made an average per 25 trials of 5.61 and an X of 4.09, a score which is barely significant. Bur. made a significant score as can be seen by inspecting Table 2. It should be noted that M and G both made better scores on this series than on Series 2 in which forms were used instead of colors and in which the cards were called and removed from the bottom of the pack.

Series 3. Cards with colors were used. The deck to be employed was shuffled at least five times and distributed evenly in a straight

line, and face down along the edge of a long table. Without having seen the cards distributed, the one subject, L, walked down the table and usually touched the back of each card as he guessed its color. After he had finished his guessing, he took up the cards in the order in which he called them and read off the true distribution to the recorder. This procedure was periodically checked by a second observer who noted all controls.

Results of Series 3. L was given 2,250 trials and scored 514 correct hits. This gave him a plus deviation of 64 above chance expectancy. The p. e. was 12.79 and the critical ratio was found to be 5.00. This result according to the usual criterion is significantly above chance expectancy; hence the results are considered significant.

Series 4. The cards were handled exactly as in Series 2. However, instead of calling only one card and removing it, the subject called five cards and removed these. The correct calls were checked after the twenty-five trials. The series was introduced as a substitute for Rhine's "DT" method because of L's refusal to attempt to guess the cards down through the whole deck without removing the cards that were called. He contended that he could not "visualize" or "imagine" the individual cards.

Results of Series 4. L was given 1,500 trials and scored 300 hits. His score was, therefore, exactly chance expectancy.

Series 5. The same type of cards were employed as in other series in which color stimuli were used. The deck to be used was shuffled at least five times, cut and placed in an even pile before the subject. The subject guessed "down through the deck" without touching the cards or removing those that were guessed. The subject continued to guess until he was told that he had guessed 25 times, or he continued until he saw that the column on the record sheet was filled since the column had places for 25 calls. When the 25 calls had been made, the subject read off the true distribution to the experimenter. The correct calls were checked and the number reported to the subject. Likewise at the end of 250 trials, the average was calculated and the subject told his score.

Results of Series 5. The subjects were carried forward through this series. M and G were each given 2,500 trials. G alone made a significant score on this series. He called 611 correctly out of 2,500 trials and he made an X of 8.24. His average score of correct hits per 25 trials was 6.11. It is interesting that his critical ratio dropped from 21.4 in Series 2 to 8.24 in Series 5. This may be due to the fact that

calling "down through the deck" seemed to be baffling and reflected adversely on motivation. The other possibility is that this series is more strictly controlled and that some unknown extraneous cues have been ruled out which were operative in Series 2. It would seem profitable to direct research toward isolating these cues.

Series 6. The procedure was exactly the same in this series as in Series 5; however, stimulus cards with forms were used.

Results of Series 6. G called 601 correct out of a total of 2,500 trials in this series. This gave him an average of 6.01 and a critical ratio of 7.50. M again failed to make a significant score and called 496 correct out of a total of 2,500 trials as compared with 518 which was his score in Series 5.

Series 7. The cards with colors were used in this series. The cards were shuffled at least five times and stacked in an even pile behind a wooden screen. This screen was built of ply wood for a top and 3/4 inch by 6 inch pine board sides. The two sides were six inches wide and eight inches high and set at an angle of 45 degrees to each other. The top projected beyond the width of the sides and extended in a semicircle. The diameter of the top was 12 inches at the widest place. When the cards were placed under this shield it was possible for the subject to manipulate them but it was impossible for him to see the cards. The subject guessed the top card, removed it by placing his finger and thumb on opposite edges, and placed it in a pile.

Two experimenters, an observer and a recorder, worked with two of the subjects. With L, only one experimenter was available. After the subject had finished calling the 25 cards of each series, the observer took the pack and called off the true distribution to the recorder. The number of hits per 25 calls was reported to the subject and his average score on two hundred and fifty trials was given him.

The two experimenters were used in Series 7 after G had made an error in reading off the true distribution. The error was apparently due to his divided attention during the checking process. He would look both at the card and at the record sheet and in so doing called once or twice when observing the symbol on the record sheet rather than the card's symbol. This error was discovered early in the series and checks of records indicated that it possibly had not previously occurred.

Results of Series 7. By inspection of Table 3 it may be seen that by applying our previously accepted criterion of significance, X of 4 or above, that M did not make a significant score but that the scores of

both G and L are significant, G having an X of 7.35 and L having an X of 5.97.

TABLE 3.

RESULTS ON 6,250 TRIALS WITH THREE SUBJECTS IN SERIES 6—1935-1936.

Number	Date	Subject	Number Trials	Number Hits	Average Per 25.	Dev.	P. E.	X
1.....	1935	G.	2500	599	5.99	99	13.49	7.35
2.....	1935	M.	2500	520	5.20	20	13.49	1.47
3.....	1936	L.	1250	307	6.14	57	9.54	5.97
Totals.....	....	....	6250	1426	5.70	176	21.33	8.25

It is to be noted that up until this time, the subjects have been oriented toward the cards either through vision or touch or both. It is, however, impossible for us to state definitely at present how cues could have been received from the particular cards of a pack. In general, the findings of Series 7 confirm the findings of Series 6. M failed to score on the more strictly controlled experiments whereas the scores of G and L continued to be sufficiently high to be called significant although somewhat reduced. This series is believed to have been very crucial for G but it would have been desirable to have had two observers during the work with L.

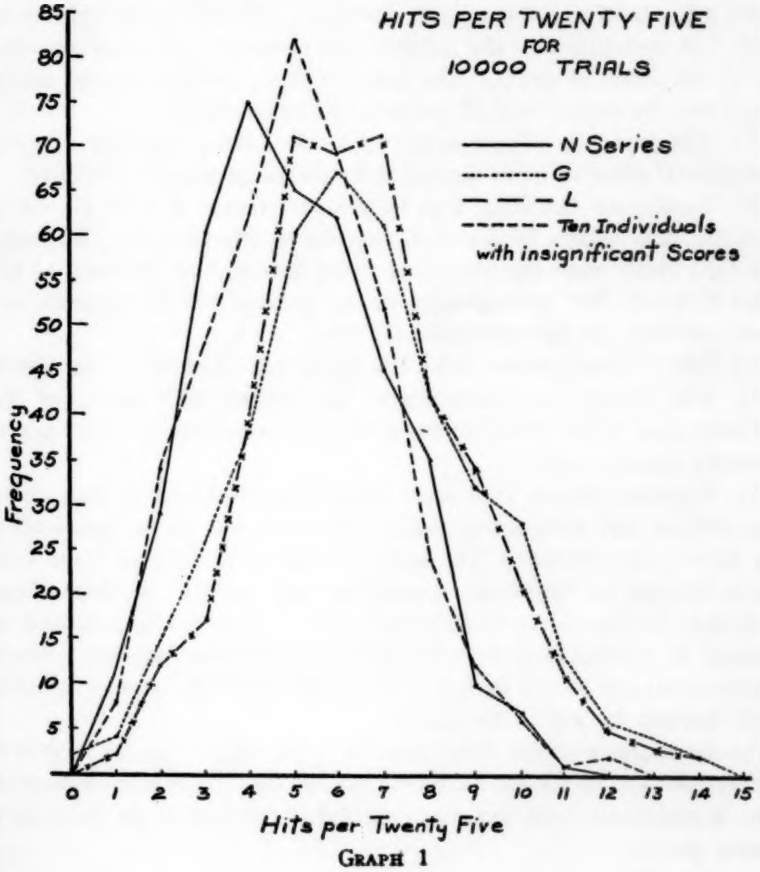
Series 8. Cards with the five colors were used, the colors were red, yellow, green, blue and white. There were two experimenters, one of whom handled the cards while the other did the recording. One of the experimenters sat opposite the subject on the other side of a table from him, and a wooden shield, rectangular in shape, made of ply wood and four feet long and three feet high separated this observer and all cards from the subject. As the experimenter removed the top card from a thoroughly shuffled deck and without noting the color, put it face down on the table with an audible click, the subject made his guess as to what the color of the card was. The second experimenter recorded the guess and so on throughout the series. When the pack was finished, the correct distribution was read off by the experimenter handling the cards, the distribution was checked against the subject's guesses, and the number of correct calls reported for the series of 25 trials. The subject, as usual, was told his average score at the end of 250 trials. This procedure was considered to be better controlled than any of the other series except possibly Series 5 and 6.

This series was carried out during the rush of commencement week and it was felt by all experimenters that G was considerably distracted.

Results of Series 8. G was the only subject used in this series. He was given 2,500 trials during which he made 472 correct calls. He made an average score of 4.72 and his deviation gave him a negative critical ratio of 2.08 which is not significantly different from chance expectancy according to our accepted criterion.

Following this series and on the same afternoon that the series was completed the experimental procedures were reverted to those used in Series 2. G was given 250 trials and he made an average score of 7.5. It is suggested that this type of control be frequently employed in this type of work.

Summary. For the purposes of this experiment the following generalizations seem to be indicated: 1. That certain of the *numerical* findings of Rhine and his collaborators can be duplicated under conditions



GRAPH 1

described by him and in general used in this paper. 2. For a subject who can guess cards with scores beyond chance expectancy, color stimuli are as effective as form stimuli. 3. Four subjects were found out of a total of sixty-five who made significant scores on one or more experimental series. However, each of the four subjects failed to score significantly on one or more series of trials.

EDITORIAL NOTES

The possibility that their data might bear upon other conclusions than those noted prompted the writers to request editorial discussion. The following point seems to us of interest:

In the report of Series 7 it is noted that the "Series is believed to have been very crucial for subject G." The hypothesis to which the results are crucial is not expressly discussed. However, the critical ratio of 7.35 indicates that the subject was responding in some effective way to the stimulus cards. The basis of this correct response can be judged only by considering all the possible explanations.

1) The presence of two observers renders the probability of gross experimental error or overt deception by the subject highly unlikely.

2) Inadequate shuffling, that is, the persistence of card groups in particular rank orders in the deck, may be readily checked. Matching each card order with the preceding order in the first 50 runs of G's Series 7, yields 243 correspondences, as against 250 theoretically expected, whereas the subject made 311 hits.

3) Direct visual contact with the cards was blocked by the screen.

4) The subject had kinaesthetic and tactual impressions of the card only *after* it had been called, a condition eliminating *direct* use of secondary sensory cues.

5) Suppose sensory cues were learned by memorizing their order of occurrence and recognizing visually the same cue on the cards while they were being checked. The cues so learned could then have been used to increase the likelihood of guessing later cards by inference from knowledge of the cards already laid aside. The learning would be indicated in gradual improvement as more identifications were made. Its inferential use would be indicated by the markedly greater number of hits toward the end of the series.

An examination of the data yields the following hit scores in groups of 250 trials: 75, 59, 63, 63, 53, 54, 62, 45, 58, 69. There is no indication of the improvement with practice expected if the results are those of a learning process.

If, however, there were learning not indicated by the above figures (as might occur if all available cues were learned on the first few runs, or if the decks of cards were frequently changed) another check is still available. When each run of 25 is divided into five consecutive groups, and these groups totaled for all 100 runs the results are (1) 120, (2) 107, (3) 133, (4) 113, (5) 126. The inferential use of learned cues, if offered as an explanation of the positive deviation of 26 in the last group, would be inadequate to explain the greater deviation in the central group, and the marked deviation in the first group, where it would be least applicable.

TELEPATHY IN THE PSYCHOPHYSICAL LABORATORY

LUCIEN WARNER and MILDRED RAIBLE

ABSTRACT: Seventeen subjects were asked to judge a large number of times whether the second of two weights presented to them was lighter or heavier than the first. A difference limen was established for each subject (the point at which the subject gave consistently three correct responses out of four trials). These weights were then presented in two ways; first, without the experimenter knowing the correct response, and second, with the experimenter knowing which weight was lighter or heavier. The number of correct responses was generally greater when the experimenter knew the weights presented than when he did not. In two cases out of the seventeen the difference was marked enough to indicate that the experimenter's knowledge influenced significantly the judgments of those subjects.

Dr. Warner and Mrs. Raible are both trained psychologists. The research was conducted in private laboratories.

WUNDT IS REPUTED to have said that if telepathy were possible all his work had been in vain. He, and most other psychophysicists, proceeded on the assumption that telepathy does not occur. So firm was this conviction that no controls for a telepathic factor were introduced in the psychophysical laboratory.

Generally speaking, however, an experimenter is expected either to control factors which might conceivably influence his results or to prove them to be non-effective. To assume the non-existence of a factor is a convenient escape but hardly a certain one.

In conducting discrimination tests the experimenter often manipulates the presentation of the stimuli and is aware of what the subject's response should be. If telepathy occurs it may be that the experimenter's knowledge influences the subject's judgments. Only occasionally has discrimination testing been done in such a way that the experimenter was unaware of the correct response until after the judgments had been made and noted. In such cases telepathy could hardly be considered a factor in the determination of the judgments. Psychologists have often failed to describe the experimental situation with

details sufficient to inform us whether the conditions were such as to give opportunity for telepathy to occur. So long as disbelief in such a possibility was so widespread that no question was raised, the experimenters were justified in their attitude. However, recent investigations, particularly those of Professor J. B. Rhine and his associates, propose the possibility of some sort of telepathic communication so insistently that the experimentalist is now faced with the task of proving its non-occurrence or of controlling for its effect in his own work. *The burden of proof is upon him.*

This report is not concerned directly with the proof or disproof of telepathy. Tentatively, it accepts its possible occurrence and proceeds with an attempt to determine the influence, if any, of this hypothetical process upon threshold determination. Weight discrimination was selected as a convenient test and the Method of Constant Stimuli was employed.[1] Seventeen subjects were tested. Both sexes were included and the age range was from six to sixty. The approximate threshold for each was first determined, and sufficient practise at this level was given to obtain reasonably uniform performance.

In each trial the subject lifted first the standard weight and then one varying from it. He reported either "Heavier" or "Lighter", referring to the relative weight of the second one lifted. The testing was done in series of 20 with pause for relaxation between series.

For convenience, it was assumed that the threshold lay at the point midway between perfect discrimination (100 per cent correct judgments) and performance based entirely upon chance (50 per cent correct). Thus the weight-difference which, for our purposes, represented the subject's threshold was that which resulted in five errors per series. This level having been established for a subject, the experiment proper was begun, using this weight-difference. In one-half of the trials the experimenter was entirely unaware of which pair of weights was being presented to the subject until after the judgment had been made and recorded. Thus a telepathic influence could not have operated in these trials, which will be called hereafter "unknown" trials. In the other half of the trials the experimenter knew which weights were being lifted and concentrated upon this knowledge during the period when the subject was making the discrimination. In these trials, called hereafter "known" trials, it is possible that telepathic influence may have had a bearing on the judgments. "Known" and "unknown" trials were given alternately, the subject not being told

[1] A description of this method may be found in most texts on experimental psychology. See, for example, Gardner Murphy, *General Psychology* (New York: Harper & Brothers, 1933), p. 228.

which was being given. Sometimes a series of twenty began with a "known," and sometimes with an "unknown" trial, this being determined by chance. The subjects were unaware of which trials were "known," but they were acquainted with the general purpose of the experiment. They were not informed of the results until after the conclusion of the day's series.

If telepathy has no bearing upon the judgments, we should expect an equal number of errors to be made on "known" and "unknown" trials. Should the results for the two kinds of trials differ reliably, we have no recourse but to attribute that difference to the fact that in one case the experimenter knew which stimuli were being presented, while in the other he did not, for in no other respects did the "known" and "unknown" trials differ.

PROCEDURE

The apparatus and procedure followed closely that customarily used in psychological laboratories. The subject sat at a table with one arm (all but one preferred the right) resting on a cushion before him, his hand reaching through a screen which shielded his hand, the weights, and the experimenter from his sight. The weights were hard rubber cylinders 1½ inches in diameter, 2½ inches high, and loaded with lead. They were borne on a turntable so they could be rotated to a point immediately beneath the fingers of the subject, who lifted them, singly, in response to a signal given by the experimenter by means of a clicker. The clicker's sound was uniform (even deliberate effort could not appreciably alter it). Thus it could not have been the source of involuntary cues. Visual cues were lacking since, as stated, the experimenter and apparatus were completely shielded from the subject's sight. The experimenter's lips were deliberately kept closed during the testing, and it seems impossible to those conducting the experiments that the "whispering hypothesis" could account for the results in any way.

The turntable, a foot in diameter, contained four depressions, in each of which was seated one of the four weights used in a given series. The entire top of the table, including the walls and floors of the four holes, was covered with felt and the weights rested on pads so they could be lifted and replaced silently. The table rotated on bearings and was practically noiseless. The four weights used in each series consisted of two standards of 100 grams each and one weight lighter and one heavier than the standard, each varying from 100 grams by the same amount. These were arranged in pairs, as in Figure 1.

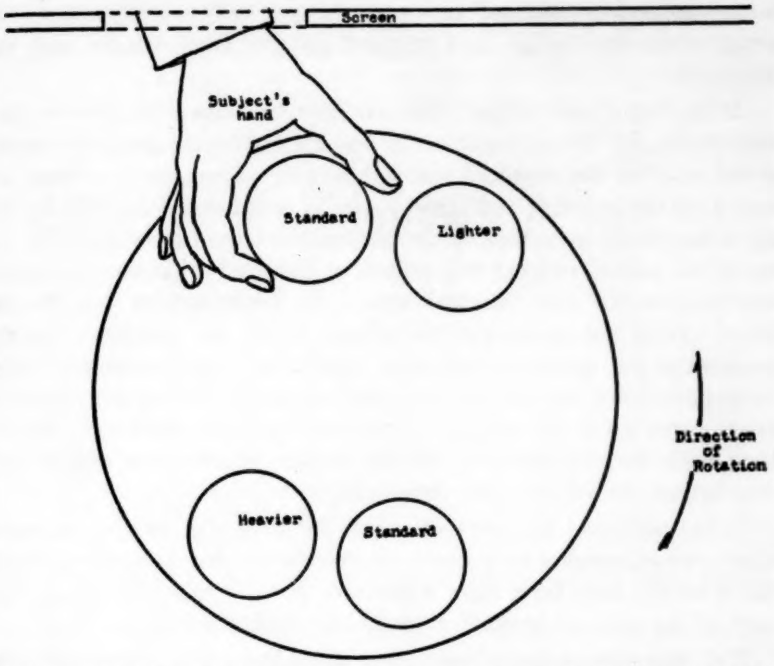


FIG. 1.—Position of weights on turntable. As shown, chance has stopped the table for presentation of the pair which should elicit the judgment, "Lighter." The hand is in position to lift the standard, after which the table will be rotated so that the lighter weight will be in position to be lifted.

In each trial the subject lifted first a standard weight and then one of the weights varying from the standard. Having done so, he reported, "Heavier," or "Lighter" in terms of the second lifted. His judgment was then recorded and the next trial given. The sequence of the pairs of weights was determined by chance by whirling the table between trials. Whichever pair of weights stopped nearer the subject's hand was presented. The experimenter adjusted the turntable so that the standard weight would be immediately beneath the subject's fingers before the signal to lift it was given. Immediately after it had been lifted and replaced, and without further signal, the experimenter rotated the second weight beneath the subject's fingers, and this was at once lifted and replaced.

When first testing a subject, the weight-difference used was well above his threshold. By degrees the weight-difference was decreased until the threshold was reached. As stated above, the experiment proper began only after the point was determined at which a subject

made reasonably uniform scores of 75 per cent correct calls. The weight-difference which had resulted in this performance was used thereafter.

In the experiment proper, "known" and "unknown" trials were given alternately. By visual inspection it was impossible for the experimenter to tell whether the standard and lighter pair of weights had been presented, or the standard and heavier, but he could determine this by feeling a thumbtack imbedded in the underside of the turntable. The tack under one pair of weights was smooth to the touch, that under the other was roughened. On "known" trials, the experimenter felt the tack before giving the signal for the subject to lift the weights. He thus knew what the subject's judgment should be. On "unknown" trials, the experimenter did not do this until after the subject had lifted the weights and made his judgment and this had been recorded. Thus in these trials the experimenter did not possess information which would have helped the subject make his decision.

It has been said that between trials the table was spun to determine which pair of weights would next be presented. It was spun so rapidly that it would have been quite impossible for the experimenter to keep track of the pairs of weights even had he wished to do so.

The following weights were used in addition to the two 100 gram standards:

80	grams	120	grams
85	"	115	"
90	"	110	"
95	"	105	"
96	"	104	"
97	"	103	"
98	"	102	"
99	"	101	"
99.5	"	100.5	"
99.75	"	100.25	"

These were calibrated within an error of .006 grams through the courtesy of Professor George H. Walden of the Department of Chemistry, Columbia University.

RESULTS

The results are summarized in the accompanying table. The columns, reading from left to right, give (1) the initials of the subject, (2) the total number of judgments made (these vary widely only because some subjects were able to work with us far longer than were others), (3) the total number of errors made, (4) the number of errors which oc-

Subject	Total Judgments	Total Errors	Errors on Known Trials		Errors on Unknown Trials		Actual Deviation of Proportion	Standard Error of the Proportion	Deviation Divided by Standard Error*
			Number	Per Cent	Number	Per Cent			
R. R.....	1900	456	237	52	219	48	2	2.3	—0.87
S. G.....	640	166	88	53	78	47	3	3.9	—0.77
C. W....	440	109	58	53	51	47	3	4.8	—0.62
G. W....	380	101	53	52.5	48	47.5	2.5	5.0	—0.5
L. S.....	240	61	31	50	30	50	0	...	0
J. E.....	280	71	36	50	35	50	0	...	0
E. W....	320	73	37	50	36	50	0	...	0
P. R.....	420	97	49	50	48	50	0	...	0
H. G....	1220	314	158	50	156	50	0	...	0
A. B.....	200	52	24	46	28	54	4	6.9	+0.58
H. F.....	200	54	29	53.7	25	46.3	3.7	6.8	+0.54
O. F.....	360	92	40	43.5	52	56.5	6.5	5.2	+1.3
E. S.....	660	154	68	44	86	56	6	4.0	+1.5
L. G.....	180	48	18	37.5	30	62.5	12.5	7.0	+1.8
W. M....	340	83	34	41	49	59	9	5.3	+1.7
M. M....	900	215	84	39	131	61	11	3.3	+3.3
L. W....	780	192	71	37	121	63	13	3.5	+3.7
Totals...	9460	2338	1115	47.7	1223	52.3	2.3	1.03	+2.23

*In this column the plus (+) signs indicate that the deviation is in the direction to be expected if there is a telepathic influence helpful to the subject; the minus (—) signs indicate deviation in the opposite direction

curred on "known" trials, (5) the percentage of the total occurring on "known" trials, (6) the number occurring on "unknown" trials, (7) the percentage of the total occurring on "unknown" trials, (8) the deviation of this obtained proportion from the 50-50 proportion which would be expected if there were no telepathic influence, (9) the Standard Error of the proportion, (10) the deviation in terms of the Standard Error.

The following formula was used in determining the reliability of the obtained deviation from expectancy:

$$\text{Standard Error of the proportion} = \sqrt{\frac{p(1-p)}{N}}$$

in which p is obtained proportion and N is the total number (of errors). This method is that most commonly used by geneticists in determining the reliability of the deviation of a given proportion of phenotypes from the proportion to be expected on the basis of an assumed Mendellian mechanism.

It will be seen that eleven of the seventeen subjects scored approximately the same number of errors on "known" and on "unknown" trials. Their deviations from the expected 50-50 ratio are all less than one Standard Error. It is noteworthy that the six subjects whose scores showed deviation from expectancy of more than one Standard Error *all deviated in the same direction.* All made fewer errors on "known" trials. The effect of the experimenter's knowledge appears on the

whole to have helped rather than hindered the subjects. In the case of two of them the deviation from expectancy is more than three times the Standard Error.

It seems to us that there are but two possible ways of explaining these deviations. First, we may presume that they are, in fact, the results of telepathic influence of the experimenter upon the subject. If that is true, we note that the influence is inoperative with most subjects. Further, that even in the case of those most benefited by it, it operates but feebly. We are led to wonder why, if the subject can "read the experimenter's mind" at all, he cannot avoid *all* errors on the "known" trials. Is it because of defects in transmission by the experimenter, or in reception by the subject? A fruitless effort to answer these and other questions by encouraging introspective comments yielded a large number of remarks made by the subject immediately after calling his judgment. These were recorded and later classified as follows:

(1) Those indicating absolute certainty, based upon the effort required to lift the weights.

(2) Those indicating uncertainty on the basis of sensory cues, but a definite "hunch" as to what judgment should be made.

(3) Those indicating complete uncertainty, the judgment being purely a guess.

Correlation of these comments with the data failed to reveal anything significant. As would be expected, comments of the first type occurred somewhat more frequently after the subject had made a correct judgment (but the difference was not great). Beyond this no relationship was noted. Type (2), which one might expect to occur on "known" trials followed by a successful judgment, occurred just as frequently on "unknown" ones.

The second, and we believe the only other possible explanation of these results, is that they reflect not the operation of the postulated function of telepathy but the operation of one of the officially recognized senses. In every case the experimenter attempted to eliminate the possibility that the subject could profit by any movements or sounds that the experimenter made. Critical consideration of the conditions leads us to believe all modalities beyond suspicion except hearing. Possibly the subject could hear the experimenter's breathing, the squeaking of his chair if he made an involuntary movement, or the scratching of his pencil as he recorded a judgment. True, it seems far fetched to suppose that such cues could in any way be construed so as to aid in the making of correct judgments. It is also true that our subjects (and there seems no reason to doubt their word) were entirely unaware of

any assistance from such sources or from any sources other than the feeling of the weights themselves. Nevertheless, the issue at stake is too important to permit even a remote possibility to remain unexamined. For this reason, further discrimination tests are being made, but not under the lax conditions usually prevailing in the psychophysical laboratory. The experimenter and the subject, instead of being separated by a mere screen, are confined in non-communicating rooms. Results will be reported as soon as possible.

SUMMARY

Seventeen subjects made weight discrimination judgments at the threshold level under two conditions: (1) When the experimenter knew at the time of the discrimination whether the response should be "Heavier" or "Lighter"; (2) when the experimenter did not know which response would be correct. In the cases of eleven of the seventeen subjects the records under the two conditions were quite similar. The remaining six made appreciably fewer errors under the former condition. For two subjects the deviation was more than three times the Standard Error.

It is not claimed that this work was conducted under ideal conditions, but rather that it was conducted under conditions usually prevailing in the psychophysical laboratory. We conclude only that under such conditions the subject is sometimes influenced by the experimenter's knowledge, though in exactly what way we cannot now say.

Alarmed, perhaps, by the sweeping nature of the problems introduced, most psychologists have avoided even mention of the word telepathy. Worse than that, they have introduced no controls against the logical possibility of its existence. On the basis of the present work, it is concluded that such controls *should* be employed as a matter of laboratory routine, regardless of whether the deviations herein reported are the result of telepathy or of the operation of one of the recognized senses in a way not now understood.

It is true that with most subjects the influence of the experimenter (however it may have been transmitted) was slight. An experimenter neglecting to consider it may err little more than the Newtonian physicist erred because of his ignorance of Einsteinian hypotheses. But one cannot forgive the refusal to recognize the logical possibility of an error's existence just because it is found that its magnitude is not great.

BENDER ON EXTRA-SENSORY AND SENSORY FORM PERCEPTION

DONALD K. ADAMS

ABSTRACT: A subject who appeared to have ability in clairvoyant perception was given cards with letters of the alphabet drawn or printed upon them and asked to describe her visual images of the symbol screened from her sight. She made an unusual number of correct responses. The development of these images was studied in relation to the responses of another set of experiments in which the same cards were presented in very dim light which was slowly increased until the symbol could be seen.

Dr. Bender's research is reported in a monograph entitled "Zum Problem der Aussersinnlichen Wahrnehmung" published by Johann A. Barth of Leipzig in 1936. The reviewer, Dr. D. K. Adams, is Assistant Professor of Psychology at Duke University.

HERE IS ANOTHER investigation of extra-sensory perception from the psychological laboratory of a university. Its author, Dr. Hans Bender, is *Assistent* in the Psychological Institute at Bonn. The experiments he reports were conducted between May, 1933, and November, 1934, with a subject, Miss D., whose putative extra-sensory capacity Bender discovered in the course of experiments on automatisms elsewhere reported (*Psychische Automatismen,* Leipzig 1936). It will be instructive at many points to compare his procedure with that of another recent investigation which has brought this field forcibly to the attention of psychologists; namely, that of Rhine (*Extra-Sensory Perception,* Boston 1934). There are many points in which his procedure closely parallels Rhine's and many others in which it differs materially.

Like Rhine, he proceeds cautiously from rather loose to stricter conditions of experiment. His subject, Miss D., was a student of philosophy "with psychological and linguistic gifts, artistic interests and capacities (music, drawing), an active sportswoman. Very differentiated, a better than average observer, and intellectually alert. Lively day dreams, including plastic images. Visual memory . . . frequent realistic dreams which are remembered in all detail. Rapid apprehension

of a personality and psychological situations. Very responsive as to mood. No marked synesthesias. . . . Good health except for migraine headaches accompanied by the usual visual disturbances. About three diopters myopic. Poor vision under dark adaptation." Although a specific investigation of the matter was deferred in order not to interfere with the clairvoyant experiments, Bender finds some evidence of eidetic imagery in his subject. He characterizes her, in Jaensch's typology, as "labile integrated (probably J_1) type," but believes this sensory integration opposed by a tendency to dissociation of the self.

In his first (and loosest) series of experiments, 27 cards 8.5 by 5 centimeters, were prepared, one with each capital letter of the alphabet in block print about 3 centimeters high, and one with a dot. The letters were applied with a "weakly phosphorescent reddish material" which projected a little from the surface of the cards; that is, as if embossed. A cellophane cover was attached to the embossed side by means of staples. Bender states that control experiments satisfactorily demonstrated that this cellophane cover provided "absolute protection" against tactile perception of the letters but gives no report of the nature of these control experiments. While the possibility is not explicitly considered, he evidently regards the staples (and in Series III the gummed paper) by which the cellophane was attached as sufficient assurance that it was impossible for the subject to insert a finger *under* the cellophane. The cellophane covered cards were placed in numbered, opaque envelopes by a third person with the printed side, right side up, toward the front of the envelope. The experimenter, Bender, shook up the 27 cards in their envelopes in a box as he received them from the third person, withdrew one and gave it to the subject. The subject knew the nature of the materials used; namely, letters of the alphabet.

The subject lay on a sofa, covered to the neck with an opaque black cloth. She was permitted to remove the card from the envelope under the cloth, to feel it and turn it as much as she liked and at the end of the trial to replace it in the envelope. The experimenter merely recorded the number of the envelope and the subject's report, without comparing them. Thirty-four "experiments" or, more properly, trials, were conducted under the conditions just described of his Series I, all on the same day. Each trial lasted about two minutes, but the subject was not restricted to one guess. More or less typical reports are the following:

8. (Letter H): "X is formed and B, again a balloon, it might also be P."

13. (Letter A): "Oblique lines, might be of K. At all events, straight lines converge somewhere. It might really be Λ, approximately."

28. (Letter L): "It had a belly, P or B, I am not sure."

32. (C): "Nothing definite. But now, quite clearly, C. It can only be C. I see it quite clearly, very clearly. That is downright fun."

Bender classifies the 34 such responses (sometimes supplemented by the subject's drawings) obtained, into five categories: Letter correct (20.59 per cent), letter wrong (14.7 per cent), form-quality of letter correct (23.53 per cent), form-quality of letter wrong (14.71 per cent), and indecisive (26.47 per cent). It should be noted, however, that these percentages are computed on the basis of a total of 34 reports. Since, as we have seen, a report might contain several guesses, the propriety of this computation is doubtful. Thus in the 4 trials, 41 letters in all were named, 11 other figures were described or drawn, and 4 trials were blank. Thus it would seem that the computation of percentages should be on a basis of a total of 52 instead of 34. If, on the other hand, it be urged that each trial must be regarded as a unit, irrespective of the number of different guesses it may contain, we are faced with the difficulty of evaluating such a trial as Number 10, which Bender assigns to the category "Figure part (form-quality) correct" on the basis of the last sentence of the report:

10. (Y): "Hard to say. At first I thought U. Then it got something above like a J with a marked upper straight cross stroke. Then it could be H. It could also become A. There come two lines which meet."

Or

15: (Q): "At first I thought: Rhombic H. *No roundings.* [Italics mine.] Now nothing more comes. But I just like to wait a minute: Now a line passes over this as if it were crossing it out. Now a circle which now and again rises like an S or a J. A bright circle on a black ground."

This trial is likewise referred by Bender to the category "Figure quality correct."

It should be said, however, that Bender explicitly attaches little weight to the statistical values given by his tables and intentionally eschews statistical analysis (to which, of course, his data—only 134 trials in all—are inadequate) in favor of detailed individual examination of each trial, its full protocol and conditions, and attempts to justify his choice in a section which we shall have later to examine.

Of more interest than these dubious percentages are the analogies which Bender, evidently a well-trained psychologist, saw between his subject's reports in this series and the reports customarily obtained in the investigation of several perceptual processes much better known than "clairvoyance." These processes are form perception in indirect

(peripheral) vision and the perception of figures in the region of the threshold of illumination. The analogies are indeed striking. They led him to orient his further inquiry toward a more precise and comprehensive comparison of the phenomenology of these processes with that of "clairvoyance."

In his second series, 27 cards of the same size as those used in Series I were prepared. Twelve of these were embossed with phosphorescent material, the remainder drawn with pastels and inks of different colors. This time the front of each card was covered with two layers of cellophane, together one mm. thick, and the back of the card with a plain card. In addition, six of the twelve phosphorescent letters and six of the others were covered, under the cellophane, with black opaque cardboard. The cards and their various covers were bound together with five staples, three at the top of the letter and two at the bottom, so that the subject could find tactually the correct position of the card. Again the cards were prepared and placed in numbered envelopes without the experimenter's knowledge as to which letters were made with which materials and which twelve were covered with opaque black cards. This was designed to assure conditions of "pure clairvoyance" uncontaminated by telepathic transfer from the experimenter. Extremely critical parapsychologists might still object that *someone* did know about the cards and the numbers corresponding to them, and that even though this someone did not know of the time or order of presentation to the subject, the possibility of telepathic transfer was not perfectly excluded. I am unable to evaluate this criticism for want of knowledge as to the conditions under which, if at all, telepathy occurs. Presumably the subject, although knowing the general nature of the materials, did not know that some of the cards had been covered with opaque black cardboard.

In the first 21 trials of the second series, which comprised 69 trials in all, the presentation of the cards to the subject was effected in the same manner as in Series I. From trial 22 on, the presentation was in three stages: Stage I was in the unopened envelope in a wooden box with the letter upright facing the subject at a distance of 70 centimeters from her, on a shelf above the sofa. The wooden box was later replaced by a sheet metal one. In Stage II, which was soon discarded, the card was in the unopened envelope, on the shelf in front of the box. In Stage III the subject was handed the envelope under the black cloth, to be opened and handled as she saw fit; i.e., the procedure in this stage was the same as in Series I. The experiments were conducted in a darkened room and nearly always, at the subject's request,

accompanied by gramophone music, which seemed to her to favor the condition of "passive concentration" which she believed most effective. As in Series I, the subject drank one or two cups of strong coffee before beginning each day's work.

The trials of Series II were conducted in eleven groups of from 2 to 13 trials each, on nine different days. The duration of a single trial varied from one to eight minutes and was usually three or four minutes. An "empirical chance" series was conducted by drawing a card at random from another box for each experimental card used. The protocols for Series II are even fuller than those for the first series. The classification and tabulation of results is subject to the same difficulties as in Series I and to even greater degree since more guesses per trial were made. The series of 69 trials yielded 15 hits; that is, according to his method of reckoning, 21.74 per cent, while the chance series yielded 2, or 2.89 per cent. The experimental series yielded 21, or 28.99 per cent of the category "letter parts (form-quality) correct," while the chance series yielded 10 cards, or 14.5 per cent in this category. This comparison with the "empirical chance" series is somewhat misleading since the subject might make several different guesses or varyingly definite characterizations and several different drawings for each trial.

In the series of 69 trials, 20 trials had to do with cards covered with black cardboard. Since in these cases the possibility of both normal visual and normal tactile perception is most clearly excluded, they are of special interest. Only one of these was correctly hit, and that one was in Stage III. I quote the protocol for that hit:

29. (E) Stage II: "Peculiar, when the envelope is put out there I have a total impression of brightness." Stage III: "I believe it is a straight letter. I thought that even before. E, perhaps. The vertical. Or B. I don't know. Something triple."

Bender himself is quite convinced that neither visual nor tactile perception of normal sort was possible in the case of any of the cards and regards this series as convincing proof that clairvoyance had occurred: "It is intended to prove only that the ability to perceive without the use of the known sense organs has been observed in the subject investigated, under the air-tight conditions described, in the conviction of the author without any doubt whatever" (p. 56).

Since, as we have seen, his data do not justify even the very slight statistical treatment (computation of percentages) to which he submits them, it is necessary to inquire as to the basis of his conviction that this series demonstrates clairvoyance. This requires examination

of the specific protocols themselves and it must be said that there are a number of very striking incidents among them, of which only a few can here be briefly described. One of these with a protocol short enough to be reproduced here was:

47. (W) Stage I: "There is something quite funny: I saw on the card for a short time the picture of a sailing ship. When the envelope is in the little box I also see the edges of the card." [At this point she made a drawing of a sailing ship in which a W was clearly discernible.]

Stage III: "Something round—no, that is wrong. That is merely waving. Now again this ship. There are oblique, upward striving lines and sharp angles." [At this point she made another drawing which was simply the W of the previous ship drawing, minus some of the accessory lines which made it a ship.]

Another striking case occurred in trial 52, in which the presented card was D. Unknown to anyone connected with the experiment, including the third person who had prepared the cards, the D had been incorrectly bound; that is, with the three staples which were supposed to characterize the top, at the bottom. I quote the protocol.

52. Stage III: "Nothing clear. Will try again. Now, fairly clearly, a perpendicular and rather to the right of the card [drawing] a sort of half circle but I don't see it right. [She had drawn a mirror image of a D; that is, a vertical at the right closing a half circle at the left.] Tell me again, please, how the position of the letter is to be determined."

The experimenter reminded her that the top of each letter was indicated by the three staples. The subject verified the position of the card by feeling over it with her thumbs. Suddenly, in the only violation of instructions of which she was guilty throughout the investigation, she threw back the opaque cloth, looked at the card, and replaced the cover. She was quite shocked by this involuntary violation of the conditions and the experimenter asked for a precise protocol. Since Bender attaches much importance to this incident, it may be desirable to quote the following in full.

I had such an uncomfortable feeling that there was something wrong that I asked how the card should lie. Then suddenly, more or less automatically, instinctively, I peeped. I didn't do it voluntarily. I was astonished at myself, caught myself at it. I noticed a D-form, but it was not a true D. Now I say to myself, after consideration, that the D must have lain wrong side to. The cellophane was wrong (the subject meant, of course, the staples). Never before have I had a desire to peep. When an impression cannot be ordered, a sort of psychical nausea occurs. If the materials had been figures I would have had no such reflex movement. But I know that I am to see letters. Since I often have the lively suspicion that on many cards there is nothing, there is something uncanny about it. My only consolation is that the D was drawn in black on white before that occurred. I often see something

that is not immediately a letter, but I can always connect it with some letter. This time I had the feeling that the impressions could not be brought together into any letter. p. 40.

While it is impossible here to give further illustrations, inspection of the very full records and the reproduced drawings, together with Bender's conviction that tactile perception had been excluded by the cellophane covers, makes understandable his belief that this series provides proof of clairvoyance. Thus where his subject gave more than one guess, the guesses are usually not wholly unrelated but often show a readily discernible direction toward the final or the most confident guess. Again in the case of the black-covered cards, there were frequent reports of "black night," "black-covered surface," and finally, the remark that she would "like to know whether there are not some cards among them on which nothing is printed." Another instance: The K card which was drawn with red ink was closed at the top by the middle staple so that visually it presented an R with an angular instead of a rounded upper portion. In the four times that this card occurred in the series of 69 trials, it was reported in trial 13 as a letter composed of straight lines and acute angles, one of the oblique lines going from lower left to upper right; in trial 14 as an R with a pointed instead of a rounded "upper balloon"; in trial 23 as certainly a clear K; and in trial 65 as follows: "Might be an E. At all events, something double. I believe see verticals. Three lines which form two parts. The vertical is important. I cannot say more. It is unfavorable to work too long on one."

His discussion of the phenomenology of these "perceptions" is again especially interesting. The subject was accustomed to project her hallucinated images upon the ceiling and the distance of the projected image became one of her most important criteria of certainty of judgment. In many experiments she awaited the rise of the images with closed eyes and when a figure survived a change in the mode of appearance, she was especially confident. She characterized the moment of opening her eyes as the most fruitful test. If the form that had appeared in the visual gray was then projected on the ceiling, she felt quite certain. Her hardest task was "to select the correct one among the numerous forms that arose" and she took pains to discard perseverations from preceding trials, various entoptic phenomena, *muscae volitantes,* etc. In this connection she found straight line figures, and especially vertical ones, much easier to distinguish than rounded ones. The "least certain" trials were those with the black-covered cards. The presence of an observer materially reduced the number of hits.

This series was followed by a series of 30 trials with sealed materials provided by the British Society for Psychical Research, including numbers, geometric figures, letters, and words. Bender attributes the negative results of this series to "an unfavorable psychological situation. The indefinitely delayed possibility of verification created unpleasantness and nervousness. During this period also the subject's general mood and bodily condition were considerably impaired" (p. 52).

In his third series new cards were prepared on which letters 1.3 centimeters high were *printed* instead of drawn and covered with cellophane fastened this time with gummed paper instead of with staples, placed as before in opaque, numbered envelopes by an assistant. A duplicate series with the same numbering was prepared for comparative experiments with gradual illumination. Both the "clairvoyance" (Stage I) and the illumination experiments were performed with the subject sitting before the projection apparatus used for the latter. On each working day the clairvoyance trials were conducted first and at their conclusion the duplicates of the cards drawn for these experiments were exposed in random order in the gradual illumination apparatus. Presentation State II in this series consisted in presenting the envelope containing the card to be extra-sensorially perceived in the unopened envelope upon the opaque cloth before the seated subject. Stage III, as before, permitted the subject to remove the card from its envelope under the cloth. The duration of the single trials, which were made in four periods on two different days, was even greater than before, and once amounted to twelve minutes. Even more guesses were made per trial in this series than before and more drawings are reproduced. Tactile perception seems in this series to have been quite satisfactorily excluded unless one supposes a quite fantastic degree of hyperaesthesia. Of the 17 trials of this series, the records of 8 contained correct guesses, which, according to Bender's method of reckoning, would be 47 per cent of hits, with 4 or, by the same reckoning, 23.5 per cent of correct part figures or form-qualities. In this series, however, 68 different letters were guessed and 28 additional drawings made or figures described so that a percentage of hits based on the total number of "guesses" would be materially smaller (8.3 per cent). Here again the question of how to treat the results arises, and the best judgment would seem to be again that any quantitative treatment, including his percentages, is misleading. His material is simply not of a kind or amount that can be so treated.[1]

[1] Thus in trial 13 (with the letter X) the subject drew K, A, V, W, X, C and G in that order as well as two additional figures. This is scored as a hit. If all

Again, however, inspection of the complete protocols and comparison of them with those of the gradual illumination experiments makes it plain that while his method of reckoning overstates his quantitative case, the alternative method of basing the percentage on a total number of "guesses" understates it. For the "guesses" are not independent but in many cases show a development quite comparable to that shown in the visual perception of the same forms under gradual illumination. There were four correct hits and six part hits in this series on Stage II presentation. In one of the four trials in which the colored letters of Series II were used, the color was given in Stage II presentation. It is only fair to say that inspection of the protocols makes a total impression more favorable to his case than do his efforts quantitatively to summarize them.

Additional subjects were used in the gradual illumination experiments and a detailed comparison is made of these perceptions with those of his clairvoyance subject in the clairvoyance experiments. Among the "errors" common to the two modes of "perception," are spatial displacements, including inversion and rotation, assimilation of the total form to its dominant form-quality (e.g., C's and G's as circles, K's as symmetrical X, etc.), duplication phenomena, and Gestalt completion.

In a final fourth series of experiments under still stricter conditions, Presentation Stage I consisted in fastening the card (in its envelope), drawn from the box at random by the experimenter, upon the back of a screen which stood behind the seated subject's back. The subject believed that the envelopes were being fastened on the side of the screen toward her. Stages II and III were as in Series III. For this series of 14 trials, Bender wisely abstains from giving his usual tables of percentages, but does point out that in 7, or 50% of the trials, "the letter was correctly named, drawn, or recognizable in the drawing." Here again the multiplicity of figures drawn or named deprives this "frequency quotient" of much meaning. He concludes his brief "quantitative evaluation of the results of Series IV" with this paragraph:

> It should be kept in mind that in ten of the 14 trials the formation of the correct letter figure was already determined with certainty in Stage I, which demonstrates the success of extra-sensory perception with the object presented behind the back of the subject.

---

repetitions, both erroneous and correct, are counted in this series the percentage of correct hits is 12.7.

This orientation toward the correct letter figure in Stage I is by no means so perceptible from the drawings to the reviewer as it is to Bender. Nevertheless, as before, one comes away from examination of the full protocol with the impression, which cannot be justified quantitatively, that the similarities of the drawings and descriptions with the "stimulus" letters are somewhat more than fortuitous.

So much for the details of Bender's investigation, which probably fail to do it justice. Careful reading of the whole report creates a general impression, which no review can substitute for, of careful, conscientious work, faithful and unusually complete reporting of conditions (physical and psychological), and alertness to the psychological bearings of both conditions and results.

Does Bender's investigation alone prove, as he believes it does, the occurrence of clairvoyance? In the writer's judgment it does not. While it is possible to agree largely with Bender's critique of the use of probability theory, the question still arises as to what convinces him that he is dealing with "extra-sensory perception;" and the answer plainly is, despite his rejection of quantitative treatment, the disproportion of correct and near-correct hits. (Query: disproportion to what, if not "p"?)[1] In fact, the only definition or criterion of "extra-sensory perception" that can be given *is* exactly, "disproportion of correct hits under described conditions excluding sensory perception." As an extreme case of this rejection of the statistical approach, there is the man (an acquaintance) who does not "believe in" Indian arrowheads. If he is shown one he will say, reasonably enough, "but, of course, with all the stones there are in the world some would be bound to look like that by chance." He can and will say the same thing if you show him a dozen or a hundred. It does not influence the sturdy skeptic's beliefs; it merely makes them patently unreasonable. In much the same way, the occurrence of extra-sensory perception can be "proven" only by piling up a large *and reliable* (in the statistical sense) enough disproportion of correct hits to make any other belief unreasonable. For Bender the proportions he found (e.g., 7 out of 14 *trials*, not guesses) are "enough." I doubt that they are enough for the majority of skeptical psychologists.

If, however, one's skepticism has already been weakened, say by wrestling in vain with Rhine's statistics, and it be conceded that Bender actually was dealing with "clairvoyance," his experiments would then have systematic psychological significance (in a region in

[1] The "empirical chance" series is not a sufficient answer to this question, since one's confidence in *it* depends ultimately upon its approximation to "p".

which Rhine's work was lacking) by virtue of their elaborate introspective reports. It is these which permit Bender to see the parallels to form perception in peripheral vision, in gradual illumination and in eidetic imagery. These analogies, of course, prove nothing and explain nothing. But if Bender actually was dealing with extra-sensory perception, they might well constitute an important step *toward* explanation and the ultimate incorporation of this hypothetical process into the body of systematized psychological knowledge. Unless his major thesis is established these parallels are merely interesting psychological curiosa.

THE TYRRELL EXPERIMENTS

BURKE M. SMITH

ABSTRACT: The central theme of this experimentation is the development by Mr. Tyrrell of a mechanical apparatus for the testing of extra-sensory perception. This apparatus has five boxes, one of which the subject is to choose and open. The task involves a more automatic than conscious choice. The experiments were made chiefly with a single subject, although others were tried. As improvements were added to the machine the subject commonly failed, but later overcame the difficulty. She was moderately successful in the few tests given with ESP cards.

Mr. Tyrrell's work was sponsored by the Society for Psychical Research (London). The reviewer, Mr. Smith, is a member of the staff of the Parapsychology Laboratory at Duke University.

AN ENGLISH INVESTIGATOR in the field of parapsychology, Mr. G. N. M. Tyrrell, has reported in a series of articles the results of experiments carried out for the purpose of demonstrating the occurrence of extra-sensory perception. These experiments were made principally with one subject, Miss G. M. Johnson, a close friend of the experimenter, whose parapsychic gifts Mr. Tyrrell had observed and reported as early as 1922.[1]

Although Miss Johnson gave great promise in the brief tests taken in 1922 no further investigations seem to have been made during the next twelve years. But in 1934 following the publication of *Extra-Sensory Perception* by J. B. Rhine, Mr. Tyrrell began a systematic study of Miss Johnson. He began by departing from the card-calling methods used in his early tests (with playing-cards) and by Rhine (with ESP cards), and adopting a technique involving a manual response. One of five small boxes was to be chosen and opened by the subject instead of reporting orally. This is the method used in the work of the second report, published in 1935.[2]

The first apparatus devised by Mr. Tyrrell for the manual technique

[1] G. N. M. Tyrrell, The case of Miss Nancy Sinclair, Jr. *Journal of the Society for Psychical Research*, XX, (1922), 294-328.

[2] G. N. M. Tyrrell, Experiments in Extra-Sensory Perception. *Journ. S. P. R.*, XXIX, (1935).

was exploratory and tentative, and possessed, as he was the first to recognize, certain defects. Of main interest here is the fact that he obtained with Miss Johnson on this pointer-apparatus scores that were well above the average expected from chance. Approximately 32 per cent of the trials were hits, while only about 20 per cent are expected by chance. This percentage held over a large number of trials, about 30,000, and compares favorably with the scoring in the results reported by Rhine.

The early pointer-apparatus consisted simply of the five boxes in a row, a screen, and a pointing rod. The experimenter, behind the screen, inserted the rod into one of the padded boxes, as nearly noiselessly as possible. The subject when signalled, opened that one of the boxes toward which she had an impulse. Her responses were rapid and relatively automatic. The experimenter's choice of the boxes was subjective, not mechanical.

The most urgent need felt by Mr. Tyrrell for the improvement of the apparatus was that of eliminating possible sound cues given by the insertion of the pointer. Accordingly, small electric bulbs were installed in the boxes, which were made light-proof. The subject succeeded in scoring above chance with the lights also. This was the first of a long series of mechanical advances reported by Mr. Tyrrell and takes us to his third paper,[8] on which this discussion will be principally based.

A complete description of the "Tyrrell Machine" as I shall call the elaborate apparatus which Mr. Tyrrell has developed, would take considerable space, and is unnecessary here. The essential features of it can be given. Since it was developed step by step and since the subject, Miss Johnson, has advanced by degrees in the "conquest" of it, I will follow the actual story in this brief review.

In place of the pointer to be inserted in the five boxes, there was now a small bulb in each hinged-lid box. A ready signal was given by an electric sounder, and a trial was automatically recorded. Then the subject opened one of the boxes, trying for the lighted one. If she opened the correct box, a success was recorded; if not, a miss. The most important advance, the critic would say, was the provision of a mechanical selector, to determine which box would be lighted. Hitherto, the choice had been made by the experimenter entirely subjectively, thus allowing a possibility of a non-random selection. The mechanical selector was believed by Mr. Tyrrell to insure random selection of the box to be lighted.

[8] G. N. M. Tyrrell, Further Research in Extra-Sensory Perception, *Proc. S. P. R.*, XLIV, (1936), 99-166.

But unfortunately the subject's scoring dropped to chance with the new machine. She said she was distracted by the "too mechanical" conditions. Mr. Tyrrell then reverted to a less mechanical method to encourage the subject. First he cut out the selector and used keys to light the lamps. Still there was failure. Going back to the use of the pointer method, however, brought back good scoring. With the pointer either thrust inside the boxes or even with it merely touching the back of the box and not visible to the subject when the correct box was opened (thus not informing her of success or failure) successful scoring went on at about the usual average of 30 to 32 per cent.

Successful again, the subject was ready, Mr. Tyrrell thought, for the lights. She was tried at them, but without the selector this time, and she succeeded, but only slightly. She was put back at the pointer again with usual success, and again at the lamps with failure. Patience and considerable regard for the subject's attitude were evidently required.

Next, the keys which were used to light the bulbs were moved from the original distance of 12 feet, to the subject's table. Then for the first time the subject succeeded with the key-lighted lamps, quite as well as she had with the pointer method. It is not stated whether the keys were ever moved back to the greater distance or not.

Then the commutator was introduced. This was a device that would automatically alter connections between keys and lamps. When the experimenter did not know the "combination" or shift made by the commutator telepathy would presumably be ruled out. Otherwise, the test would be, as in the pointer method, one of general or undifferentiated ESP. After a period of failure with this novelty good scoring was again achieved and another advance made.

Another innovation was made in the method, and one that is unique in this field of experimentation. This is what Tyrrell calls the "delay-action" technique. The peculiarity of the technique lies in the fact that the circuit is so closed that the lamp in one of the boxes will be lit only when and if the correct box is opened. This condition was accepted immediately and successful results were obtained.

Again, the Tyrrell machine offered the possibility of testing for precognitive clairvoyance, or what seemed to Mr. Tyrrell to be precognition. With the ready signal given to the subject a half second before the key was pressed, the choice was made before the lamp was lit. The commutator was used with a view to excluding telepathy. The 2,255 trials yielded 23.9 per cent of hits, and this is very significant, giving odds of 10^5 to 1 against a theory of chance.

In all there are reported 47,876 trials by Miss Johnson with the Tyrrell machine, and the average of 23.6 per cent gives odds of 10^{80} to 1 against chance.

If there is no question that something beyond chance is required to account for the results, what other alternatives can be reasonably considered? First there is the question of sensory cues. The precognitive tests at least ruled these out of consideration, since the response was made before the key was pressed.

Next arises the question of whether a truly random distribution occurs in the experimenter's choice. If not, it would be possible for similar peculiarities of distribution to occur, due to non-ESP factors, and yet producing extra-chance scores. The machine as it stands needs the further improvement of a recording device to register the actual boxes lit and opened. From these records then could be checked the personal peculiarities of both experimenter and subject (agent and percipient).

As a matter of fact, the small sample records given in the report do show peculiar distributions with Mr. Tyrrell, Miss Johnson, and a third subject. The point of greatest concern is that all three emphasize the middle box outstandingly. This is a serious matter, although the samples are small. In any case the question should be cleared up experimentally. Has it been done already?

The use of the commutator would bear on this question, but while shifting the experimenter's distribution of choices it would not alter it enough to make it a truly random selection. Personal patterns would still occur with it. Only the mechanical selector would escape this possibility, and Miss Johnson failed to score above chance with that device in operation. However, Tyrrell reported a long series of 3,261 trials at an average of 21.7 per cent, and odds of 125 to one, carried out using the mechanical selector to determine the choice and using the keys, with choice unknown to the operator. These odds (125 to one) are taken as significant by Tyrrell, though they are somewhat below the standard used in the Duke experiments and other similar work. These conditions exclude explanation in terms of similarity of peculiar personal patterns between experimenter and subject. The "precognition" tests and "delay-action" trials are not clear then of this alternative hypothesis.

Bearing upon this point of personal patterns and peculiarities is the discovery of Mr. G. W. Fisk, a co-worker of Mr. Tyrrell. He found that if he knew the success or failure of each trial, he could score by system. He need only avoid a box for a time after a hit

occurred. This was due, he found, to Mr. Tyrrell's tendency to avoid duplication of boxes in his selection. The crucial question would seem to be, "Did Miss Johnson share this peculiarity with him?" It would be only natural to do so, since it requires training to avoid the tendency. This point, too could be checked from the data if complete records of the experiments were kept.

Mr. Fisk's demonstration that extra-chance scoring can be done by his system using the ESP cards, *provided* he knows immediately when he makes a hit, has no bearing on work done by others with ESP cards, since the rule is to inform the subject of success only at the end of a run. Mr. Tyrrell's card work, which I am about to mention, is also free from this danger.

Turning now to the ESP card tests given by Mr. Tyrrell, we find the numbers of trials very limited. The investigator felt urgently the need to give the subject a highly automatized response method with little conscious cognition involved. The few card tests were therefore only for comparative purposes.

To continue first with the principal subject, Miss Johnson, there were first 1,000 trials with mixed conditions, yielding results giving odds of 150 to one against chance and then 1,000 rigidly controlled trials, with screen and two observers, yielding scores giving odds of 75 to one against chance. According to Tyrrell's standard this too is beyond the criterion of significance. The cards were not looked at, and the conditions provided a test of clairvoyant perception. There seems no alternative hypothesis applicable to this card data, unless the criterion of significance be regarded as too low.

Mr. Tyrrell gave ESP tests to other subjects, using both the pointer method and ESP cards. The number of subjects is not large because he found work with Miss Johnson of primary importance. He tested 29 other subjects with the pointer method, and 20 subjects of the same group with the ESP cards. One other successful subject is reported for the lamp apparatus, although it is not clear how many were tried.

In the 37,100 trials made with the pointer apparatus with the 29 other subjects, three made significant scores. In the 4,600 trials made with ESP cards with 20 subjects besides Miss Johnson, only one subject reached significant scoring. But proportional to the number of trials, this would be equivalent to eight as compared to the three with the pointer. The number of trials is too small for comparison, however, as Mr. Tyrrell points out. What is important is that both methods succeed and that more techniques are available for the work.

By these tests with other subjects Mr. Tyrrell greatly strengthens

his case in relieving the work with Miss Johnson of sole responsibility for the conclusions. It is perhaps somewhat less likely that the other subjects possessed the same habit patterns (in choosing the boxes) as Mr. Tyrrell. Certainly in the ESP card tests, which yielded two significantly successful scorers in 21 (including Miss Johnson), no such consideration can properly be raised.

Mr. Tyrrell concludes, and probably most readers will agree with him, that he has demonstrated ESP. Some will be more impressed by the fewer, but better controlled card tests, others by the machine or the pointer apparatus. For those who are bothered by the question of normal similarity of habits of selection, Mr. Tyrrell points to the instances of failure. Failure, he thinks, would not be so likely to occur if scoring depended on similar habit-patterns of selection between experimenter and subject, as it would with ESP itself. There is too the just significant case of Miss Johnson's work with the choices made by the mechanical selector, but converted to the key method by the experimenter. Fortunately, perhaps, so much evidence from so many sources has now accumulated that one may regard such results with more tolerance. Also Mr. Tyrrell's work is still going on, and presumably will be strengthened on just these points. The further perfection of the machine will help, provided the subjects succeed with the mechanical selector, and provided it can be perfected to give a well distributed "random" selection.

As it stands now the selector used would appear, from the figures given in the report, to depart significantly from a random distribution. There was a 7.5 per cent slighting of one box through 2,000 control trials. This gave only 370 selections for the box in question, or 30 below the expected 400. This should be corrected and doubtless will be. The perfection of this machine for all the purposes aimed at by Mr. Tyrrell, and its successful use in demonstrating ESP, is so high and so important a goal that we must await with patience the necessarily long climb that is involved.

If ESP has been demonstrated in the Tyrrell work, the other conclusions are of still greater importance, particularly the very tentative one concerning precognition. All the more interest therefore attaches to his further improvement of the tests. Mr. Tyrrell's emphasis on the role of related physical and mental factors in ESP is most appropriate. He points to the need of health, of favorable emotional states, of interest, and even of subconscious processes favoring this delicate ESP capacity. He is strongly inclined to favor the use of the machine and

offers some very good reasons. The question, however, is one for experiment.

Of course, experimenters may themselves have preferences but the relative value should be tested in actual practice. Does a more motor (and less conscious) mode of choice-response or judgment favor success more than one involving conscious choice, as in card calling? It is to be hoped Mr. Tyrrell will go further with comparison of the machine with card tests, and settle this point which he has raised.[4] In any case, there are some advantages clearly recognizable in the mechanical instrument, enough to warrant its fullest development. Mr. Tyrrell is unquestionably making a very important contribution to the field of parapsychology.

[4] Editorial note: This problem was taken up by Woodruff and George following the work of Mr. Tyrrell. See their paper in this number of THE JOURNAL.

SOME BASIC EXPERIMENTS IN EXTRA-SENSORY PERCEPTION: A BACKGROUND

EDITORIAL NOTE: None of the articles of this number of The Journal provides an historical background, and all therefore assume some acquaintance with the earlier work on the part of the reader. As was indicated in the Editorial Introduction this extract from the most recent report of research in this field is offered to help orient the reader to some degree toward the experiments described in this issue. Footnote references will guide him to further literature of importance.

This extract is taken from an article entitled "Some Selected Experiments in Extra-Sensory Perception" published by J. B. Rhine in the "Journal of Abnormal and Social Psychology" for September, 1936.*

EXTRA-SENSORY perception, under the headings of "telepathy" and "clairvoyance",[1] is not new to science, nor to its psychological branch. Many psychologists have given active attention to it,[2] and some have contributed to the evidence for its occurrence. Among the latter are Bechterew,[3] Brugmanns,[4] Coover,[5] Estabrooks,[6] and Janet.[7] Of these, Coover apparently did not recognize his own contributions at the time, but several critics have since agreed that his results are positive, while

* Reprinted here with permission of the editor.

[1] Telepathy is the perception of another's mental state by other than the known sensory means; and clairvoyance, perception of objects without sensory action. "Extra-sensory perception" includes both.

[2] For a more extensive review see J. B. Rhine, *Extra-Sensory Perception*, (Boston: Bruce Humphries, 1934), Chap. 2.

[3] *Zeitschrift fur Psycho-Therapie*, VIII (1924), 280-304.

[4] "L'etat passif d'un telepathe controle par le phenomene psychogalvanique," *Comptes Rendu du IIme Congres International de Metapsychique*, (Warsaw: 1923).

[5] Experiments in Psychical Research, *Psychical Research Monographs*, (Stanford: 1917), No. 1.

[6] *A Contribution to Experimental Telepathy*, (Boston Society for Psychic Research, *Bulletin V*, 1927).

[7] For selected instances and references see Richet, *Thirty Years of Psychical Research*, (London: Macmillan and Company, 1922); and *History of Psychology in Autobiography*, ed. Murchison (1930), Vol. I, p. 125.

Janet has simply remained silent regarding his findings, allowing others to interpret them. The Danish psychologist, Lehmann, who, with Hansen, offered the only serious criticism of the early telepathic studies, later came himself to be convinced of its actual occurrence.[8]

The most extensive study of telepathy and clairvoyance yet made was begun at Duke University in 1930. The first two years of experimentation were described in a brief report entitled "Extra-Sensory Perception of the Clairvoyant Type",[9] and the first three years were covered by a monograph, "Extra-Sensory Perception", published in 1934.[10] Another report gives the results obtained with a special subject, a professional medium.[11]

The reports mentioned submit the results of well over 100,000 trials made in clairvoyant card-calling and in telepathic thought-transmission. The experiments in which the better conditions obtained seem to allow no other interpretation of the results, and appear to demonstrate an extra-sensory mode of perception. Evidence is offered not only of telepathic and clairvoyant perception, but of the natural character of these capacities. That they have some relationship to certain known functions is shown by the study, and a close interrelationship between the two, telepathy and clairvoyance, is indicated, if not established.

The purpose of this article is to give a brief account of some selected experiments in which both good conditions prevailed and good scores resulted, along with a full account of conditions maintained, and a report of controls and safeguards used. The selections made are chosen to meet the various possible questions or alternative hypotheses that we have to consider in this work. These fall under the following headings: For the clairvoyance tests, the question of sensory cues on the cards used, or of poor shuffling; for the telepathy work, the question of similar habits of order of selection made by the agent and the percipient in choosing and calling the symbols used. And, finally, for both clairvoyant and telepathic perception tests, the question of whether chance alone can explain the results obtained. While these results are selected from our present stock of several hundred thousand trials, they are taken fairly, including all the consecutive trials made under the conditions stated. A given block of data is taken intact, without any omission of results within the series. There was no "selection" in the improper sense of the word.

[8] After the above was written, the report by Dr. Bender of his experiment in clairvoyant perception appeared. See Dr. D. K. Adams' article on Bender's work in this issue, p. 63.

[9] *Journal of Abnormal and Social Psychology*, XXIX, No. 2, (1934), 151-171.

[10] Rhine, *op. cit.*

[11] *Character and Personality*, III, No. 2, (1934), 91-111.

GENERAL METHODS

The tests for clairvoyance were all made with cards hereinafter called "ESP Cards". These were used in packs of twenty-five, each pack containing five each of five simple designs: ★ + ≋ □ O (star, plus, waves, rectangle, and circle). The cards were cut from white heavy cardboard which was opaque to light from a 100-watt light bulb. The figures were stamped in fountain pen ink with rubber stamps which left no detectable warping or other tactual stimuli.

In the tests for telepathy, the same symbols were used, but, to exclude the possibility of clairvoyance, the mental images of them alone were used without any objective representation until after the test when the record was made. This made the test one of pure telepathy, with "mind-to-mind" transference alone possible.

The results were scored by runs of 25. The scores of a series of runs are totaled, and the mean chance expectation for that number of trials found by taking 1/5 of the total trials made in the series. This mean value (np) for 300 trials (12 runs of 25) would be 60. Let us suppose the total score is 90 (an average of 7.5 hits per 25). Then there is a positive deviation from the mean (np) of 30. This is then measured by means of the probable error, which is found by taking $.67449\sqrt{n\,p\,q}$, which here would be $.67449\sqrt{300\times 1/5\times 4/5} = \pm 4.7$. The ratio D/p.e. (*i.e.*, the ratio of deviation to probable error) gives a value that I have been calling X, which can be converted into a fraction of probability P by means of Normal Probability Integral Tables. In the illustration given we would have $X = 30/4.7 = 6.4$, which would give a P of about 1/50,000; that is, the odds against chance as an explanation of the results are 50,000 to 1, which would of course amount to practical certainty on the point.

This method is similar to that used by Coover and Estabrooks and is a standard technique in the field.

THE NEW CARD SERIES

In this series with subject HP, I was myself the recorder and observer. Twenty-five packs of freshly printed cards were used in the series, and the subject had never seen them before. The packs were laid on the table in a stack at the beginning of a sitting, and each was used for three runs, with the exception of two that were used only once and five that were used only twice. A pack was handed to the subject by the observer, and he was asked to shuffle it. The shuffling was always thorough. Then I cut the pack and placed it before the subject, or allowed him to hold it face downward in his hand, removing

each card still face downward to another pile on the table after he called it. No one looked at the faces of the cards until the end of the run. There had been no opportunity for anyone to learn the order of the cards nor to study them for sensory cues. The cards were checked after the completion of each run of 25.

In 1650 trials consisting of the first, second and third runs with these new cards, the average per 25 trials was 9.5. This is nearly double the mean value and the X value is very high, 27.0. To give the odds against this being due to chance would require a number of 70 digits. It will be recalled that an X value of 4 is the usual criterion of significance. In these scores the first runs with the new packs were as good as the third.

The results may be grouped as scores in hits per 25 for the 1st, 2nd and 3rd runs as follows: (It will be seen that while all 25 packs were used for a first run only, 23 were used for the second, and 18 for the third. This was due to interruption of the experiments.) 1st runs: 4, 8, 10, 7, 12, 11, 12, 15, 8, 7, 12, 11, 13, 9, 9, 8, 11, 13, 13, 7, 6, 9, 4, 9, 7. 2nd runs: 6, 8, 9, 8, 11, 11, 12, 7, 8, 12, 7, 5, 10, 12, 14, 11, 9, 13, 8, 15, 6, 9, 15. 3rd runs: 8, 14, 10, 8, 12, 7, 5, 11, 8, 10, 6, 12, 10, 8, 8, 10, 10, 8.

The 1st runs average 9.4 hits per 25, and the 2nd, 9.8, and the third, 9.2 per 25. Each group even taken alone is highly significant in its deviations. The ratios D/p.e., or X values, are respectively, 16.3, 17.2, and 13.1.

Any one still skeptical about this mathematical procedure may get a "commonsense" picture of the results by observing that whereas in a chance distribution the scores ought to be distributed so as to give an average closely approximating 5, in these results, we have 9.5; and whereas the mode in the normal case would be 5, in this there is a bi-modal distribution, the principal mode being 8, and the second 12. And the median here is 9 instead of 5. Of the total of 66 runs, only 2 are below the expected mean and 62 are above it.

This New Card Series is submitted as evidence against the hypothesis that the results may be due to poor shuffling of the cards, leaving an order that may be already known to the percipient. It matters not how the cards were shuffled, since in the 25 first-run group, the order was entirely unknown to the subject. He had never seen any of the 25 packs before.

Incidentally, this New Card Series answers the question of whether possible sensory cues could have been given by slight markings on the

backs or edges of the cards. It would require experience with them to learn these.

THE DT SERIES

This DT (or "down through") series consists of 1,625 trials, or 65 runs, made with subject HP (myself again the observer) under the following conditions: The pack of cards would be shuffled, frequently by the subject himself under the surveillance of the observer (myself again in this series). The shuffling by HP was really profuse,[12] since he felt that he got some advantage from contact with the cards. He habitually did not look at the faces, and could not possibly have "stacked" (*i.e.*, arranged the order of) a pack with the observer watching. The observer always cut the pack and it was left lying on the table while the calls were made. About a dozen different packs were used in the series; sometimes they were new to the subject and sometimes previously used packs were provided. When detectable markings appeared on the cards they were discarded.

The peculiar feature of DT work is the fact that the pack was not touched until after the 25 calls were made. After they were all made and recorded the call-record was checked against the order of the cards in the pack.

The last five calls in the run, representing the bottom of the pack as it lay, were the best. The hits were next best for the first five at the top, and lowest for the central 15 cards of the pack, as they have been for several other subjects—though there is one striking exception.

The 1,625 trials gave a total score[13] of 482, which has a positive deviation of 157 ± 10.09. This gives a critical ratio (X) of 14 and a P of 10^{-20}.

The DT data add weight against the hypotheses of sensory cues, memory, and bad shuffling. They bear too upon physical hypotheses involving hyperaesthesia and special radiation, on which other data given below will bear. In a word, they offer difficulty to a wave hypothesis, since waves from 25 cards in a pack must give only splotchy confusion.

CAMPUS DISTANCE SERIES IN CLAIRVOYANCE

In this series are reported all the tests given subject HP with the cards in one building on the campus and himself in another. The series

[12] However, if only one card is displaced in the pack by the shuffling, and the cut is made by the observer, there is no way by which the so-called "shuffling error" can apply to the work.

[13] The individual scores are as follows, in consecutive order: 8, 8, 12, 6, 5, 9, 8, 5, 8, 8, 10, 9, 5, 5, 8, 7, 8, 6, 8, 7, 8, 9, 9, 1, 7, 11, 11, 5, 8, 4, 8, 9, 4, 4, 11, 6, 10, 11, 7, 6, 11, 12, 5, 4, 11, 8, 2, 9, 11, 7, 8, 7, 7, 8, 4, 4, 10, 7, 6, 7, 6, 8, 8, 5, 8.

is complete for these conditions. In it there are four groups and two conditions of distance. Dr. J. G. Pratt, then a graduate assistant in psychology, was the principal observer throughout the series.

The observer and subject synchronized their watches, and arranged to work at a stated time and distance. At the specified time the observer would take the top card from a shuffled pack of E S P cards in the room agreed on and lay it face down on a book in the center of the table without looking at its face. Thirty seconds later the subject in his cubicle in the Duke Library would record a call for the card. At the end of the minute, the observer would remove the card and take the next one. The cards as removed would be kept in order for later recording. Two runs were made per day.

Groups A, C, and D were made with the observer in the Physics Building and the subject in the Library. The distance was about 100 yards. In group B, the observer was farther away with the cards, approximately 250 yards. The subject was in the same place in the library as in A, C, and D.

Group E is included for comparison. It was taken with the subject and the cards in the same room, with Dr. Pratt as observer. In a part of the group, the observer held the cards, and in part the subject did. The conditions were not especially good, but the comparison with the results of the distance tests is very interesting.

In Groups A-D, the records were sealed up after each sitting and delivered to me before subject and observer got together. In Group D, I was present with Dr. Pratt as a second observer.

The cards were shuffled between runs, and this shuffling occurred just before using, with HP absent from the room. Even if only one card were displaced at an unknown point and a cut made, the shuffling would baffle any guessing even from a perfect memory of any previous order of the cards in the packs used. The only check needed in the shuffling is that obtained by correlating successive card records for the runs. For Group A this averages 4.5 hits per 25, or slightly below the mean chance expectation. Two packs were used at a sitting. Now if we take only the most favorable combinations in each sitting we raise the average only to 5.6. So the results obtained are clearly not explainable by inadequate shuffling.

The scores[14] may be represented in averages per 25 for the group, and the significance of the deviations represented by the X values.

[14] The actual scores are as follows:
Group A: 3, 8, 5, 9, 10, 12, 11, 11, 12, 13, 13, 12.
Group B: 12, 10, 6, 4, 10, 10, 2, 6, 5, 12, 7, 5, 12, 11, 9, 10, 6, 3, 0, 13, 10, 12, 12, 4, 4, 1, 4, 4, 4, 7, 6, 5, 0, 6, 3, 11, 9, 9, 4, 8, 6, 0, 6.

Group	*Conditions*	*Number Trials*	*Average per 25*	*X-value*
A	Physics Bldg. and Library.....	300	9.9	12.6
B	Medical Bldg. and Library.....	1,100	6.7	8.4
C	Physics Bldg. and Library.....	300	7.2	5.6
D	Physics Bldg. and Library.....	150	9.3	7.9
E	Same room.....	900	8.2	13.7
	Total.....	2,750	7.7	21.0

With all sensory cues excluded by the conditions, with the shuffling checked satisfactorily, and with the chance hypothesis so clearly ruled out by the statistical evaluation of the data, there is left only the hypothesis of some extra-sensory mode of perception. Anyone suspecting observer and subject of collusion must include me also in the conspiracy in view of Group D in which I was a witness of Dr. Pratt's handling of the cards.

O-Z TELEPATHY SERIES

In this series subject GZ was the percipient and Miss Sara Ownbey was both observer and agent or sender. Both were graduate assistants in Psychology, and it was felt that an independent observer was unnecessary. In Group A of this series the agent and percipient were in the same room; in Group B, in adjoining rooms, and out of sight of each other. A door was open between and about 10 or 12 feet of distance separated the two subjects. In Group C, there were 28-30 feet between and two walls, in line of vision, though again doors were open between rooms. For all three groups there was an electric fan going, which would effectually obscure any "involuntary whispering", if it occurred. Signalling was effected by a regular and monotonous tap of a telegraph key.[15]

The agent, using the symbols of the E S P Cards, but not the cards themselves, and deliberately choosing her order of symbols five in advance so as to avoid falling into habitual preferences, would think intently of her first image and then give the signal to the percipient. The percipient would call out his choice and the agent would record the symbol chosen and check it if correct. She was practised in avoiding

Group C: 5, 4, 11, 8, 4, 9, 9, 8, 9, 10, 2, 7.

Group D: 12, 3, 10, 11, 10, 10.

Group E: 10, 12, 11, 11, 9, 7, 5, 8, 8, 9, 10, 5, 8, 7, 6, 4, 8, 7, 8, 7, 6, 11, 12, 13, 12, 9, 11, 5, 9, 5, 6, 10, 4, 6, 7, 5.

[15] If it be objected that the sounds might be varied unconsciously and serve as cues, it might be replied that no opportunity to learn the cues was provided. The subject did not know when he was right or wrong.

following any routine selection, and varied her order systematically. There have been several checks made as to her success in avoiding characteristic patterns which might be followed by the percipient. These were made by correlating the consecutive records of the agent. The check results of three such series, one of which is reported in full below (the Junaluska Series) have yielded for 650 trials an average per 25 of exactly 5. This is the mean itself. Very evidently the agent is successfully avoiding repeating herself or falling into any stereotyped order of selection. A day by day check between agent's series for 8 days gave only 3.9 hits per 25.

The question of similarity in habits of choice may be further checked by cross-checking the agent's record with the percipient's of—let us say—the day preceding or following, that is, correlating runs not intended to match. A cross-check was made of the Junaluska Series given below and it yielded an average of only 4 hits per 25. Habitual similarity of order of choices between agent and percipient can, I think, be said to have been satisfactorily eliminated. Moreover, in the series cross-checked as just described, the runs were made daily and the cross-check, therefore, covers also the question of daily routines or patterns of order of choosing.

Finally, as an extreme test of "patterning" or order-habits, we may check the percipient's records against themselves in consecutive order. The percipient might have a pattern of order without its being coincident with a similar one supposed for the agent. But if *he* does *not* have such a pattern, the question of similar order-habits is ruled out. Such a check made on 12 of the percipient's records given below, taken consecutively from a block chosen at random, and on the 8 records of the Junaluska records, yields an average of coincidences of 4.6 between runs, as against an expected 5.

In this series Miss Ownbey was solely responsible for the entire records. In others she was not, as in the Junaluska Series below. In one such short series with GZ[16] as subject the average score is 11.0 and the X value is 7.5. In another, $X = 4.2$. In both, a recorder was present with the subject. The purpose in presenting this series[17] is,

[16] It should be said that GZ does not like to keep a record himself, since it distracts him; he feels that he cannot get into the right mental condition and continue to record. He sits with closed eyes and in deep abstraction from surroundings.

[17] The question has been raised as to whether the agent might not, through her strong interest in getting good results, be likely to mistake unclear enunciation by the percipient who is off in another room, and to give favorable interpretations of calls not clear. First, the names of the symbols are phonetically quite easily distinguishable; each has a different vowel sound. Second, audition was good with

however, not primarily to offer evidence of telepathy as such, but to bear on the relation of distance to scoring rate in considering the wave hypothesis of extra-sensory perception. The rate does not fall off with distance as both the wave theory and the hyperaesthesia hypothesis would require. The succeeding series satisfies the point of independent recording.

In presenting the data of this O-Z series, I will give first the 12 consecutive runs from which the check on the percipients' records was made. These were made in one sitting, with three conditions, given below. And below these results will be given all the telepathy tests given GZ by the observer under these conditions (one room and two rooms away) during the summer of 1933, and along with them all of the tests made at the same sittings with both in the same room.

	Same Room		*One Room Away 10-12 feet*		*Two Rooms Away 28-30 feet*	
	Trials	*Score Avg. per 25*	*Trials*	*Score Avg. per 25*	*Trials*	*Score Avg. per 25*
12 consec. runs..	50	15.0	125	19.0	125	16.8
Totals.....	950	14.0	750	14.6	250	16.0

These scores are so high that their significance is obvious without any computation. The 12 runs given above yielded the following scores, in order, 14, 16, 17, 21, 19, 19, 19, 13, 15, 16, 17, 23. In one of the runs of the 950 trials there was a perfect score of 25.

THE JUNALUSKA SERIES

In this series MFT was percipient, with Miss Ownbey again as both observer and agent. The tests were made with synchronized watches corrected from Western Union time. The distance was over 250 miles, MFT being at Lake Junaluska, N. C., and the observer at Durham. One run was made daily, at first at the rate of one call per five minutes and later at the rate of one every three. Reports were to be sent to me by agent and percipient immediately. As it happened the first three days' report was mailed to the observer instead of to me, but she had already given me her records, and brought the percipient's report direct from the post-office (and it was in the familiar handwriting of the percipient). The records of the next five days were sent to me. The average for these eight runs was 10.1, and gave an X value of 10.8, which is, of course, very highly significant.

---

open doors between rooms. Third, the scores for this work with the agent and percipient in the same room are nearly as high as with the two separated.

"Chance" is excluded as an explanation. Similarity of preferences or order patterns as a possible factor is shown to be lacking. The data meet all counter-hypotheses quite satisfactorily, I think, all that have ever been raised, except that of possible deliberate collusion. Those not knowing the assistant and subject may, by the very pressure of doubt evoked by such unusual phenomena, raise a question as to this. There never is, however, in any research, an absolute answer to this question. The relative assurances that can be had are briefly the following: As already stated, the observer was a Graduate Assistant in Psychology with no slightest reflection cast on her honesty during the four years I have known her. In one who is seeking a psychological training and career we do not expect to find the motivation necessary to practice a planned course of deception such as this. And the marks of deception are lacking. First, when MFT was tested in the same room with the agent-observer, just before the distance series, she scored an average of only 7.7 hits per 25 trials. At the time she did not expect to do better with distance. Second, several times before and since this series, we have most urgently wanted high scores in special experiments, in some of which the same percipient was used,—and have failed to get them. In the many series of telepathy tests that were statistically significant and were made under good conditions, more than a dozen different agents have been used and nine different subjects. There were fifteen different individuals involved, and in addition a number of observers. The collusion theory would not be seriously applied to all these by many critics. In fact, it has been raised only by those in psychic research who are close to the subject of fraud in mediumistic phenomena.

The hypotheses (and the criticisms) of "inadequate shuffling" of cards have been met by the New Card Series. That of "possible sensory cues" on the cards was likewise excluded by this series, as well as by the DT ("calling down through the pack") Series, and the Campus Distance Series. In telepathy, the question of natural habitual similarity of patterns in agent's and percipient's choices is answered by proper checks on the point in question. Hyperaesthesia, involuntary whispering, and the like are excluded by the Distance Series. "Chance" as an explanation is well ruled out by the tremendously significant deviations, and the odds against chance to which they correspond. These cover the criticisms thus far raised against the work, except that of possible collusion. It is clear that this must involve myself as well as many others, since in one of the experiments I was in a position to check this point (Campus Distance Series, Group D). It is freely granted that for the extreme skeptic this may rightly be regarded as the weakest

point. All scientific work suffers alike in this respect at some point in its development.

If we are correct in thinking that these results under the conditions stated indicate an extra-sensory mode of perception—that is, perception without the recognized sensory processes being essentially involved—no one will dispute their tremendous importance for the general science of mental life. I need not here indulge in speculation as to the bearing it would have. Some cautious suggestions appear in the monograph referred to above.[18] If these are sound, they open up such large possibilities for the study of mind and its place in nature that we need to go all the more slowly to determine the evidential grounds upon which they rest.

[18] *Extra-Sensory Perception.*