The Journal *of*
Parapsychology

Volume 10 JUNE, 1946 Number 2

EDITORIALS

CONFIRMATORY EXPERIMENTS IN PK RESEARCH

THERE is always great interest in the amount of confirmation a new finding receives whenever the discovery is important enough to justify critical examination. The ESP research is a good illustration. But the investigation of that ability has by this time been so extensively confirmed that among those who are well informed on the matter there is little interest left in mere addition to the evidence as such. In the period of more than sixty years since experimental work in ESP began, independent repetitions have numbered into the hundreds, with a large number of these based on amply safeguarded procedures.

In contrast, the newer PK investigation has been before the public only since 1943 and there has been very much less time for the spread of repetitive experiment. Nevertheless, enough progress has been made to warrant a brief summary of the amount of confirmatory work that has been published to date.

We must keep in mind, however, that there are various kinds of confirmation, with different values attaching to each variety. First and simplest, there is the kind of confirmation obtained from repeated experimentation by the same investigator. While such retrial does have importance, it is for the general reader the least valuable form. Then there is the repetition by different experimenters working in the same laboratory. Even if no other variable

is introduced this type is of considerably greater value. Higher still in the scale is confirmatory research by different experimenters working in the same laboratory but under varying procedures and conditions. The highest type of confirmation is the favorable repetition of the essential experiment by workers in other laboratories, whether under similar or modified conditions—the more independent the experimenter and his method, the better, of course, the substantiation.

The special reassurance which the more independent type of confirmation provides is due to the control it offers over personal factors that may operate in the research. The competence, reliability, and good judgment of the experimenter are all in question if a new and strange conclusion is reached. The more the demonstration is repeated by others, the less room there is left to doubt its genuineness.

Even in the short time during which PK reports have been appearing in the JOURNAL, there have been confirmations of PK representing all of the types listed above. Numerous experiments have been carried out by different experimenters in the Parapsychology Laboratory at Duke where, thus far, the greatest concentration of PK research has been localized; but fourteen other confirmatory experiments have already been reported that were conducted elsewhere. In some of these there was a certain amount of cooperative action between the experimenter and the Duke Laboratory in designing the test procedure; in others there was none. Among the completely independent series might be mentioned the Thouless coin-spinning experiment, the Gibson Machine Series, both Frick experiments, and the McDougall series.

Three reports have been published that fail to give significant results: one in the JOURNAL OF PARAPSYCHOLOGY by Nash (his first paper) was suggestive but inconclusive, and two in the *Proceedings of the Society for Psychical Research*, one by Parsons and the other by Hyde, were entirely insignificant. Nash began his work in ignorance of the contemporary research at Duke (which had not yet been published). Parsons and Hyde worked completely independently of the Parapsychology Laboratory. On the whole, then, the score of confirmatory researches in PK is a high one, however

we look at the value of the successes and failures. More independent experiments are known to be in progress in this country and in England, and reports will presumably be published in due course. Two of them have, in fact, already been described in correspondence as conclusively confirming the case for PK.

In the discussion of repetitions of the PK experiments there is often some question as to what valuation is to be placed upon an experiment which fails to confirm—one which gives only chance results. It goes without saying that no amount of *failure* to find an elusive phenomenon can prove anything except that under the conditions of the test and with the subjects who participated, it did not appear to a significant degree. This result does not tell us much, if anything, about the value of reports in which the results *were* significant. Unless some explanation can be found as to why, in one experiment, evidence of PK is obtained, and why, in another, it is not, no appreciable contribution to knowledge can be salvaged from the research giving chance results.

However, as every student of parapsychology knows, many of the experiments yielding only chance score totals have proved to be fruitful in other respects. Analyses of different character than those initially applied have in some cases revealed hidden relationships that were at first overlooked. Some of the most important discoveries concerning PK, like those concerning ESP, have emerged in this way. Experimenters, therefore, are now more hesitant to draw conclusions immediately from an experiment that shows only a chance rate of scoring.

There is a type of confirmation not yet mentioned in this discussion, though it is by now a familiar one. This is a form of reinforcement that is still more independent and objective than any of those listed. It is that which derives from the discovery of lawful relations within the data of an experiment. These relations are of extraordinary evidential value when they follow a predominantly consistent pattern through a large number of independent researches. And they suffer nothing from having been wholly unanticipated when the experiments were being conducted. The QD's (quarter distributions of hits on the record sheet) offer the best example of this type of confirmatory analysis, though not the only one. As we

have pointed out before, this evidence of consistent internal patterning has all the virtues of independent confirmation together with that rare and much sought-for virtue of being based on an operation (the analysis of the data) that is as objective as an independent experiment and is repeatable at will.

A great deal more confirmation of the PK effect may be forthcoming from a variety of independent sources, but it will be interesting to see if the evidential quality of the confirmatory support given by the QD analyses can be greatly improved upon.

J. B. R.

ESP AND PK AS "PSI PHENOMENA"

THERE is a need for a general term to include the phenomena of both ESP and PK. It is more than a matter of mere convenience, however, for there is a general impression among those who are familiar with the evidence of ESP and PK that in these different manifestations we are dealing with a common underlying process. It seems likely that the cognitive effect obtained in ESP tests and the kinetic one obtained in PK work are results of the operation of the same basic psychophysical principle.

It has been argued before that in ESP some sort of interaction of the percipient's mind with the object perceived may reasonably be supposed to occur, though it is of course unobserved. As it is, the experimental situation is designed to bring out only the *cognitive* effect and either to ignore or to prevent any possibility of an effect upon the object itself. The PK experiments, on the other hand, are designed to focus attention entirely upon the *kinetic* results produced. Yet we recognize that the intelligent direction of the falling dice (or discs) seems to require the role of a cognitive function that could not be anything but ESP. It is likely therefore that we are measuring a single basic ESP-PK reaction that merely gives the different results because of the different test procedures. Such, at least, is our present working hypothesis.

What term is available to characterize this relationship? Drs. Thouless and Wiesner, in their article appearing in the present issue,

extend their term "psi" to cover both PK and ESP phenomena and presumably to refer to whatever is common to the two. This term was introduced at the suggestion of Wiesner by Thouless in his Presidential Address appearing in the *Proceedings of the Society for Psychical Research* for July, 1942; but at first its application was limited to ESP, including precognition. Though we hesitated earlier to add further to the already long list of names applying to that section of parapsychical phenomena,[1] we do recognize the need for the term in the more general application now proposed. It may be that as our knowledge of the nature of the processes expands, a more descriptive term than the name of a Greek letter can be found, but for the time being it will be difficult to provide a more suitable designation than "psi" for the process producing ESP and PK effects. These effects are, then, "psi phenomena."

The mere acceptance of a new name, however, can be passed over too easily. What is of greatest importance here is the fact that we now can accept as a matter of course the synthesis of thought that goes with the adoption of the term. Back of this acceptance lies a large accumulation of experimental evidence which has given us much the same picture of the conditions affecting performance in PK tests as it has for success in ESP experiments. Psychological, physiological, and social influences have had similar effects upon both types of ability. Position effects are much alike for both, and both are essentially unconscious functions. The general experimental results have thus made it easy for the logical interrelation represented by a common name to find acceptance among experimenters and students of the research.

Beyond question this unifying conception of the research adds value and strength to the individual findings that are a part of the synthesis; for understanding consists in great part of just such interrelationships and the further tying-in of these findings with the more familiar processes and properties of the normal personality.

J. B. R.

[1] See editorial, "A Proposed Basis for Choosing Terms in Parapsychology," in this JOURNAL, September, 1945.

APPROACHING A NEW STAGE OF ESP RESEARCH

From the editor's vantage-point, a development of the first importance in parapsychology can be seen to have been shaping up in recent years. We refer to the convergence of a number of successful studies dealing with the traits and states of personality associated with successful performance in psi tests, more particularly those of ESP. This discussion is especially precipitated by the paper by Dr. Humphrey in the present issue, showing that those properties of the subject's personality which affect the form of the drawings he makes in ESP tests are associated with success in ESP. Even while we are waiting for the discovery of more about the nature of the traits concerned, we can appreciate the importance of the fact that the trail leading in that direction has been found.

This discovery by Dr. Humphrey recalls the work of a few years ago by Dr. Stuart in which he found that the subject's degree of affectability by the knowledge of his former score was reliably related to his success in ESP scoring. It also brings to mind the more recent researches of Dr. Schmeidler, which have brought out conclusively the correlation between favorable attitude toward ESP and success in ESP tests. A number of other findings, too, might be listed from the literature if we allow for the consideration of more general mental states found to have a bearing upon ESP success.

From a survey of what has been published and of reports now on the way to publication (which will add considerably to the picture) it is safe to predict that we are rapidly approaching a new stage with regard to the relation of personality and ESP performance. It seems likely that the more or less blindly groping phase of that research is about to come to a close, and that we are ready for a more sharply focused attack upon the problem. Enough material will shortly have accumulated to warrant a concerted effort at a planned test-construction approach in finding out in advance who is going to be a successful subject in ESP tests, and who is not.

From the scattered findings that have already accumulated, it ought before long to be possible to initiate the development of a series of trial inventories and to conduct the researches for the proper testing of their efficacy. Whatever the degree of initial success of this procedure, it will mark the introduction of an advanced order of ESP research and represent a further step toward control over the capacity involved.

J. B. R.

SUCCESS IN ESP AS RELATED TO FORM OF RESPONSE DRAWINGS:

I. CLAIRVOYANCE EXPERIMENTS

By Betty M. Humphrey[1]

ABSTRACT: This report is the first of a series, and one that introduces a new line of research in ESP. In the work reported here Dr. Humphrey divided response drawings obtained by Dr. C. E. Stuart in his ESP tests into two categories, as based upon her judgment of their form qualities or type of expression. In this paper she presents the method and also the first section of her research—namely, that part which deals with clairvoyance tests in which these drawings were the responses. She found that on the basis of her classification of the drawings as either "expansive" or "compressive" she could effectively separate the high- and low-scoring subjects. Although in the original tests Dr. Stuart had not found the subjects' efforts at reproducing the target drawings successful to the extent of being significant, Dr. Humphrey found a significant difference between the drawings rated as expansive and those rated as compressive; that is, on the basis of form quality alone she had singled out the high-scoring and low-scoring subjects from an insignificant collection.

An even more striking effect was found in connection with displacement. Dr. Stuart had reported finding that the subjects had scored significantly, not on the stimulus drawing aimed at but on the one just preceding. When Dr. Humphrey divided the response drawings into the expansive and compressive groups, it was found that the compressive group obtained a more highly significant displacement score than had been found by Dr. Stuart in the collection as a whole. Both in the displacement effect and in the direct hits on the target, the classification of the drawings into expansive and compressive successfully differentiated between high and low scorers. No claim is made as yet as to the traits of personality represented in the expressive qualities of the drawings, but there is no doubt that these objective criteria do identify real characteristics of the individuals who produced them.—Ed.

Introduction

THE PARAPSYCHOLOGIST, in his search for a greater understanding of the nature of extrasensory perception, is ever looking for conditions or characteristics, permanent or temporary, which are peculiarly associated with success and failure in the various tests for ESP ability.

[1] This article is based upon a thesis entitled "Discrimination Between High- and Low-scoring Subjects in ESP Tests on the Basis of the Form Quality of Their Response Drawings," which was submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Graduate School of

A relatively large number of researches have dealt with the problem of discovering the optimal experimental conditions for the occurrence of ESP. The subject-experimenter relationship, rewards, boredom, frustration, the length of run, and drugs (to mention but a few factors) have all been found to affect the scoring rate of the subjects in ESP tests.

On the other hand, the personality of the high- or low-scoring subject has received relatively little attention in the form of researches specially designed for that purpose. The relationship between ESP scores and intelligence test ratings was dealt with in one study in which significant positive correlations were found in two of the five ESP series analyzed (**9**). Another exploratory research into the relationship between ESP scores and more permanent traits as revealed by personality questionnaires gave only small correlation coefficients, which were not significant (**10**).

But may there not be other aspects of the personality which, if they can be revealed, may be correlated with high- or low-scoring levels in ESP tests? For instance, is there anything in the behavior of the "good" subjects which differentiates them from the "poor" ones? In short, may there not be differences in *ways* of responding in the ESP test which are associated with different levels of scoring?

If this question were affirmatively answered, we would not only achieve an added degree of control over the phenomenon of ESP, but also we would presumably gain increased insight into its nature.

Two essentials appear to be needed in order to answer this question. The first is a test which allows for a wide variety of response. We need a test in which the subject has the greatest freedom for expressing his personality. The second essential required is a method for evaluating any behavioral differences that appear. We must be able to estimate the strength or degree of the responses which are to be correlated with success in the ESP test.

Arts and Sciences of Duke University, April, 1946. The second article based on this thesis will present the GESP series and will be published in a later issue of this JOURNAL.

I am greatly indebted to Dr. J. B. Rhine and Dr. C. E. Stuart for invaluable guidance and help in this undertaking. The assistance of Miss Elizabeth McMahan in checking the data and the help of Mrs. D. H. Pope in judging the clairvoyance series are also gratefully acknowledged.

The first requirement seems to be best met by the ESP test in which the subjects make drawings in the attempt to reproduce or identify concealed pictures known as "targets" or "stimuli." In this type of test, which is one of the oldest in ESP research, the subject is usually asked to try to reproduce a drawing that is being looked at by another person, the "agent," or a drawing that is concealed in an envelope, depending upon whether the test is meant to be one of general extrasensory perception (GESP) or of clairvoyance.

The most recent formulation of the subject's task has been that by Stuart (**16, 17, 18**) in what he calls the "free response" method. He has pointed out that there may be responses other than mere pictorial reproduction which may be equally valid identifications of the stimuli. For example, a good response to the meaning or content of a picture might well be an association to which the picture gives rise, rather than a mere duplication of the elements or even the meaning of the stimulus. Thus Stuart asks the subject to orient his attention to the stimulus drawing and then to put down anything that comes to mind.

Stuart's instructions allow for a presumably greater freedom of expression than is possible in other ESP test situations. Here, then, in the free response data we may logically expect any individual differences to show up most plainly.

Having chosen the free response method as providing the most suitable material for the study of behavioral differences, we are next confronted with the need for a method of evaluating such differences. For many years psychologists have studied drawings with a view to discovering more about the personality of the subject (**4, 5, 15**). Recent researches using projective techniques have shown that, by an analysis of the *form* of a drawing—*how* it was drawn rather than *what* was drawn—it is possible to discriminate between adjusted and maladjusted individuals (**7, 11, 19**). The expressive movements employed in drawing leave tell-tale clues which enable the investigator to tell, for instance, whether the artist worked tensely and rigidly or freely and uninhibitedly. These and other signs in the drawing may be diagnosed by the trained clinician who

is thus enabled to deduce the behavior of the person who made the picture.

In the search for some method of analyzing the free response data of ESP tests, this development in clinical psychology was considered to be applicable, since the materials to be analyzed in each field are quite similar. On the strength of the clinical work, it is a reasonable assumption to suppose that the free response in an ESP test with drawings is a creative expression revealing traits of personality of the subject, which are thus available to analysis. *In the present investigation, therefore, a search was made to determine whether differences in drawing qualities are associated with differences in scoring levels in the ESP tests.*

Although a number of studies of drawings have been reported, few give criteria for using form qualities in any diagnostic way. The only fully reported set of criteria for assessing the form qualities in spontaneous drawings is that of Dr. Paula Elkisch (7). She has defined four criteria for this purpose: (1) rhythm-rule, (2) complexity-simplexity, (3) expansion-compression, and (4) integration-disintegration.

A number of sets of four drawings each from the Stuart series were examined according to the Elkisch procedure to see which of these qualities were present. The subjects were divided into two groups according to the dominant quality observed in their drawings. For example, all drawings were examined to see if they were "expansive" or "compressive," and the subjects were separated into two groups accordingly. Then the ESP scores for each group were totaled and compared. This same procedure was followed for each of the other three criteria. It became evident at once that one of the Elkisch criteria, that of expansion-compression, effectively divided the subjects on the basis of their ESP scores into high-scoring and low-scoring groups and that this was the only one that did so. Therefore it was decided to apply this one criterion as far as possible through all available ESP series in which free material was used.

The present article will present the results of the form analysis of all the available clairvoyance series, while a second report will give the results of the analysis of all available series of general extra-

sensory perception (GESP) in which an agent looked at the stimulus pictures.

The Analysis of Form

The method of analyzing form qualities in drawings has been described by Elkisch in her monograph, "Children's Drawings in a Projective Technique" (7). In addition to publishing definitions of the criteria she has used, Elkisch also gives a simple scoring procedure for each.

Elkisch's criteria were formulated after study of children's drawings and they center about the concepts of "adjustment" and "maladjustment." The actual scoring procedure gives a certain score for a quality "indicative of good adjustment," and another for one indicative of poor adjustment. In applying this technique to drawings of ESP tests we must be careful not to adopt the concept of adjustment wholesale into a situation where it is not yet known to be applicable.

It is not necessary to recapitulate the development and statistics of the Elkisch criteria, since in this study we are not concerned with their validity in clinical usage. If the one criterion of expansion-compression successfully discriminates between different scoring levels in ESP tests, its efficacy or applicability in other quarters is of secondary importance to the present research.

On page 13 of her monograph, Elkisch defines the criterion of expansion-compression as follows:

> *Expansion* is expressed (a) through the widening (opening) of the space at the drawer's disposal, by presenting only a part of the object—this might also be an abstract design—which has to be completed by imagination; (b) through the creation of a "spacious" background which may be presented on a sheet of paper of any size; (c) through the creation of an experience of space by the means of rhythm and integration. For example, a well formed representation of an explosion expands the space by "bursting" it. Such an expression conveys controlled aggressiveness, willful and forceful activity; (this is a rather virile form of expression). Expansion stimulates the imagination dynamically. It conveys an atmosphere of freedom, courage, adventure, and may be a symptom of vitality and of healthily developed extraversion.
>
> Expansion stands for a direction toward the surrounding world; for the potential ability of making contact.

Compression is based on a meticulous, fearful concept of space, expressed either in the spatial appearance of the object itself, or in its spatial relationship to other objects, or to the space at the drawer's disposal. Compression conveys a feeling of discomfort, of being shut in, of pressure and compulsion. Compression may be, if connected with other traits, a symptom of a neurotically developed introversion, even of a compulsion-neurosis.

Compression stands for isolation.

With these definitions well in mind, the judge considers each drawing of a given subject and records the quality present. Then, according to the Elkisch method, if expansion is recognizably present, a score of 1 is given. On the other hand, if compression is judged to be shown in the subject's drawing, a score of 0 is recorded. When the extremes are not obvious or easily decided, a score of .5 is given. The total score for a given set of four drawings consists of the sum of the scores for each picture. Since each subject makes only four drawings in the free response method for ESP testing, the total expansion-compression score for each set of four drawings will range from 4 to 0.

In the present study it was decided that the interval between scores of 1.5 and 2 afforded the best dividing line for the separation of the subjects into two groups according to the recognizable presence of expansion or compression in a set of drawings. Subjects with a score of 2 or above were put in one group and those with scores of less than 2 were put in another. A score of less than 2 indicates the definite absence of expansion (or the presence of compression), while a score of 2 or above indicates that expansion is either definitely present or not definitely absent.

At first these definitions seem obscure and subjective, but in actual practice they are fairly easy to apply once the main ideas have been grasped. This is facilitated by study of the examples of the different qualities which are reproduced in the Elkisch monograph. Several examples from the Stuart material may serve to illustrate the drawing qualities of expansion and compression, and the reader may see for himself the differences in the two different types of drawings. In Figure 1 two examples of "compressive" drawings by different subjects are reproduced together with two examples of "expansive" drawings.

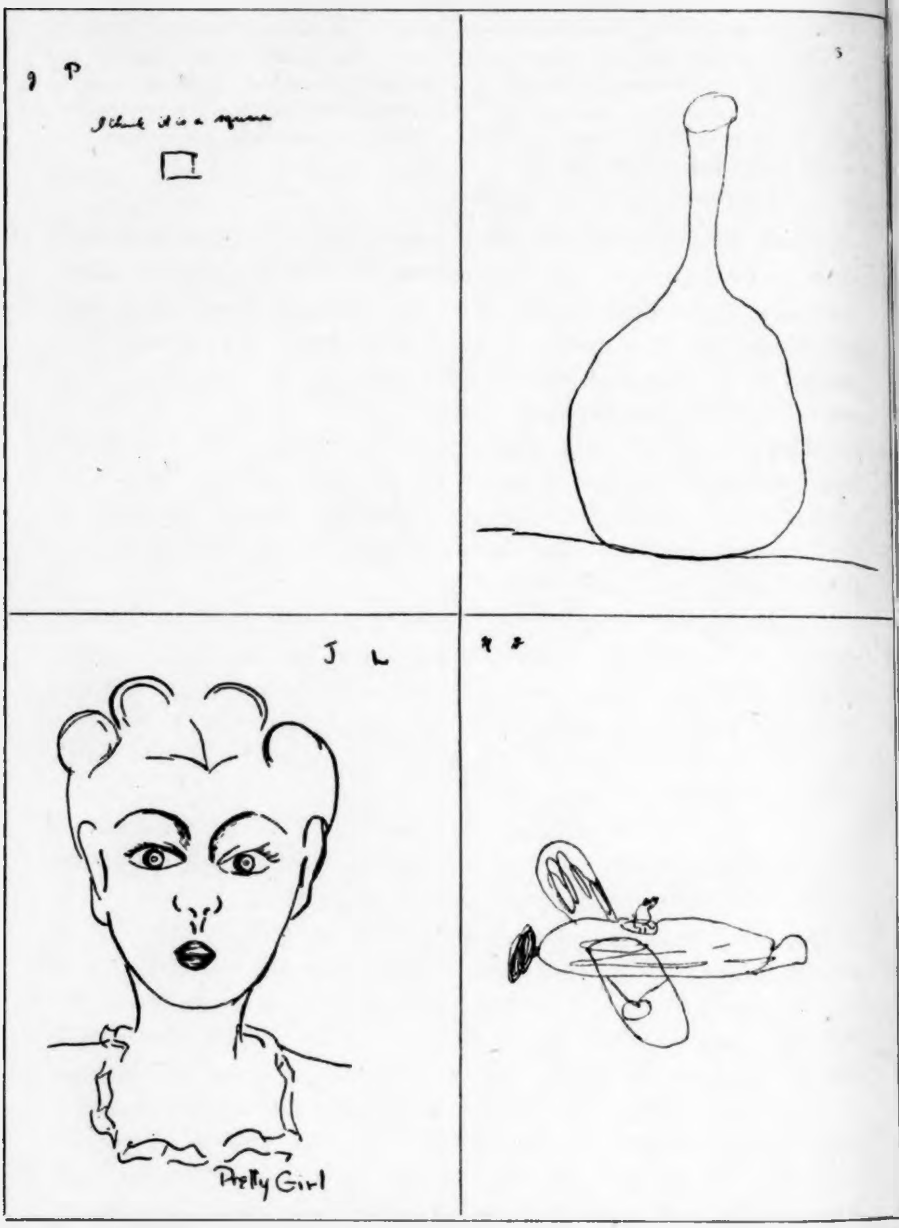


FIG. 1. Examples of expansive and compressive drawings from the Stuart series. The upper two reproductions were rated as compressive, while the lower two were given expansive ratings. (These are scaled reductions of the originals which appeared on 7½" x 10½" sheets of paper.)

The first subject's response was rated as compressive because of the timid, unimaginative use of the drawing space. The timidity of this subject is further demonstrated by the remark which she has made in connection with the picture. She has not boldly drawn a square and let it go at that. She lightly drew a few lines, hardly finishing one side of the square. In spite of the fact that she has been asked to draw, she included the uncertain statement, not that "it is a square," but that "I think it is a square." The remaining pictures in the set by this subject which are not shown here are quite similar both in drawing quality and in accompanying remarks.

In order to show that mere *smallness* of the object portrayed in relation to the drawing space is not synonymous with compression, a drawing by another subject is presented. The vase (upper right) is too large for the space and furthermore, it could hardly be said to convey "an atmosphere of freedom, courage, and adventure." On another response sheet by the same subject, we find simply the inscription: "not inspired." This is of course an indication of complete inhibition and lack of imagination. These two "responses" primarily account for this subject's compressive score.

The third drawing shown in Figure 1 (lower left) received a rating of expansion. The drawing has been skillfully executed with some feeling for line and proportion, and, moreover, the subject has not hidden behind any stereotyped, conventional design, but has expressed ideas with some degree of originality.

To show that expansion is not synonymous with drawing skill, the response of another "expansive" subject is reproduced. Here, in spite of the lack of fully developed drawing skill, the subject has expressed her idea fairly boldly. There is no indication of compression—no sign of withdrawal or "isolation." Some of the pictures by this subject were not rated as completely expansive, but as "undecided." The total score, however, was large enough to put this subject in the expansive group—that is, there was no evidence of compression seen in her responses.

These definitions by Elkisch are not without support from other studies. In their surveys of artistic behavior in the abnormal, Anastasi and Foley (4, 5) mention many of the characteristics of

"maladjusted" subjects' drawings which would come under the heading of compression according to the Elkisch system. To mention only a few: Pfister found "small, faint line, among fearful neurotics as well as chronic schizophrenics and the final stages of catatonia" (5, p. 193); Reitman reported that "the chief characteristics of the reproduced drawings by schizophrenics were their stiffness and lack of expressiveness" (5, p. 196); one of Traube's conclusions after a study of children's drawings was that "extremely small drawings were often associated with feelings of inferiority" (5, p. 219). Other examples dealing with other aspects of this one criterion could be cited, but these instances may serve to show that there is other evidence supporting the placing of small drawings, for example, under the heading of compression.

Outside the realm of drawings we have Allport's description (2, p. 414) of "expansion" and "reclusion" as being the easiest of overt expressive traits to identify:

> Our first meeting with a person gives the clue, and subsequent experience more often than not merely confirms our first judgments. . . . A person who is expansive projects himself into his social relationships; he talks readily, expresses his opinions frequently, and leaves little doubt as to his views on any subject. The reclusive person finds little to say; he relates his opinions briefly or not at all. He is reticent. Since these traits are by definition overt in expression, it is not surprising that several studies show them to be the most reliably rated of all the common traits of personality. . . . A garrulous person seems almost always garrulous, and a "human clam" is under almost any circumstances hard to pry open.

This description of observed behavior is very close indeed to that of Elkisch in regard to the "crystallized gestures" of drawings.

Description of the Series Analyzed

There have been four clairvoyance series in which the free response method was used. Stuart has given a detailed account of these in his original report (18), but the essentials may be briefly reviewed here. Stuart prepared four line drawings for each series and, after stapling each one between sheets of cardboard, he enclosed them in sealed opaque envelopes. Each envelope containing a stimulus picture was then enclosed in a second opaque envelope, and a response sheet was clipped to the outside.

Stuart then gave a set of four such envelopes to each of four student experimenters who worked directly with the subjects. (These experimenters did not know what the stimuli were.) An experimenter would give the subject one of the envelopes and ask him to draw on the response sheet anything that came to mind in regard to the enclosed picture. After the subject had drawn all he wanted to for one picture, he was handed another envelope and asked to do the same for that one. This procedure was continued until all four stimulus pictures had been dealt with.

In each of the series done by the four experimenters there were twenty-four subjects. Thus all the available clairvoyance data consist of four responses from each of 96 subjects. The subjects were undergraduate students at Stanford University. They were entirely unfamiliar with ESP tests and could hardly have foreseen the drawing analysis used here.

After each series was completed, the responses and stimuli were matched by the preferential matching method (**16**). By means of this procedure a "score" is obtained for the subjects' performance. In general, the matching procedure was carried out by the student experimenter and by Dr. Stuart, neither of whom knew the true correspondences at the time.

> The matching phase is carried out in the following manner. The judges are asked independently to rank the four stimuli with respect to the degree to which they are associated with a given response. The four stimuli are placed in front of the judge in random order. He is given one of the responses and is asked to designate which stimulus is most like the response in terms of form, meaning, association, or any criterion he wishes. Then he designates which is a second best match to the response in hand, and so on. His choices are given numerical weights: a "4" indicates that he thinks a given stimulus matches a response best; a "3" indicates his second choice; a "2" his next choice; and "1" indicates the poorest match. After this has been done for each of the responses, the process is reversed: the four responses are ranked with respect to the degree to which they are associated with a given stimulus. The four responses are placed in front of the judge and he is asked to rate each one of the stimuli in terms of their correspondence to the four responses. Weights of 4, 3, 2, and 1 are assigned to these choices as before.

In the present series, after the preferential judgments of the two experimenters had been obtained for all the responses of a given series, the true correspondences were obtained from the lists of the two judges. Stuart had retained until this time the list of the stimuli enclosed in given envelopes, while the student experimenter had held the list giving the order of the subjects' responses to given envelopes.

From the preference weights assigned to each stimulus and each response, a "score" for each matching was obtained by adding the weights assigned each correct correspondence. If each response had been indicated as matching best the stimulus for which it was actually intended, the sum of the weights assigned would be 16, a "perfect" score. If each response had been designated as being the poorest match for the stimulus aimed at, the sum of the weights would be 4, or the worst possible score. Chance expectation would be 10 for the four pictures. The total score for the subject is the total of the preferential weights assigned the stimuli and responses by two judges. Thus chance expectation for the two matchings of each judge would be 20.00, while that for both judges—the "score" for a set of four drawings—would be 40.00. The score used throughout the present report consists in the sum of the weights given the true correspondences in the two matchings by two independent judges. The evaluation of the scores obtained in this manner has been described in detail elsewhere (16, 17).

Evaluative Procedure and Methods

Since I was the one who judged all of the response drawings in terms of expansion or compression, it is important to note that at no time did I know the ESP score of any of the responses I was analyzing. I was not acquainted with any of the subjects since all of the series had been carried out at Stanford University. After all of the drawings in a given series had been scored according to the expansion-compression criterion, Dr. Stuart and I worked together to determine the results of the analysis. As I read to him the name, initials, or identifying number for each subject together with the subject's form score, he noted down the ESP scores for each. The ESP scores for the subjects with a score of 2 or above for expansion-compression were added together as one group, while

the ESP scores for subjects with a "compressive" score were pooled in another. Let it be stated unreservedly that after I had access to the ESP scores, no expansion-compression score was ever changed!

The basis of comparison in use throughout this study is the *mean score* of the sum of the weights assigned to the similarities between stimuli and responses per set of drawings. *Direct matching data* is the term applied to the mean scores of the sum of the weights assigned (per set of drawings) in the comparison of stimuli and the responses *intended* for them.

Sometimes, however, it has been found that subjects seem to identify not the target aimed at but the one adjacent to it. This "displacement" effect can be evaluated in the same way that "true correspondences" are evaluated. Here, however, instead of comparing Stimulus 1 with Response 1, one compares Stimulus 2 with Response 1, or Stimulus 1 with Response 2. If, for example, we are testing the hypothesis that the subjects more frequently identified the just-previous stimulus, we would note the weights given when Response 2 was compared with Stimulus 1, Response 3 with Stimulus 2, and Response 4 with Stimulus 3. The addition of these weights gives a mean score for the *displacement data.*

Throughout this study "Student's" *t*-test (see **13**) has been used as the test of significance applied to means and to differences between means. For interpreting the probability of *t*, "Student's" *t* distribution was used instead of the normal probability integral tables.[2]

All checking and computations connected with this research have been carried out independently by at least two different persons. This includes the complete rechecking and recomputation of all ESP scores for all series.

[2] When the number of degrees of freedom was greater than thirty, the tables of the normal distribution were used with a slight correction of the standard deviation. The correction used throughout this study when more than thirty and less than one hundred degrees of freedom obtained was discussed by Rider (**13**, p. 89). The *t* was corrected by multiplying it by $\sqrt{\frac{n-2}{n}}$ where *n* equals the number of degrees of freedom. The tables of the probability integral (for example, **12**, pp. 2-8) were then entered for the corresponding probability. When *n* was greater than one hundred, the probability of *t* was obtained without correction from the normal tables.

Results

The results of the division of the subjects into an expansive and a compressive group will be presented for the direct matching data as well as for the displacement data.

Direct Matching Data

Of the total of 96 subjects 41 were rated as "expansive" (that is, they were given a score of 2 or more in the form judgment), and 55 were rated as "compressive" (that is, they were given a score of less than 2 on the judgment of form quality). The mean score on the ESP test for the expansive group is 41.88, while that for the compressive group is 37.45, as compared with a chance expectation of 40.00. The t of the difference between the mean scores of the two groups is 3.00, which has a probability of .0013. Since there was no hypothesis as to the direction of this difference, it is appropriate to double the probability. The corrected figure would be .0026, which is still significant since it exceeds the criterion of the 1% level of significance.

Table 1

Total Direct Matching Data as Divided by the Expansion-Compression Criterion

Series	Expansive Subjects			Compressive Subjects			Difference Between Mean Scores of the Two Groups t_d	2P
	N*	Mean Score on ESP Test	t	N*	Mean Score on ESP Test	t		
W.....	15	41.73	1.06	9	36.00	2.84	2.41	.025
B.....	10	41.00	.35	14	34.21	2.62	1.91	.071
I.....	2	45.00	5.00	22	39.41	.44	1.19	.243
Z.....	14	42.21	1.03	10	39.00	.46	1.02	.285
Total...	41	41.88	1.64	55	37.45	2.71	3.00	.0026

*N refers to the number of subjects each of whom made one set of four drawings.

The statistics for the total results, together with those for each of the series separately, are presented in Table 1. There it may be seen that the trend noted for the total is found to be consistently present throughout the individual series. In each series the mean score of the expansive subjects is higher than chance, while that of

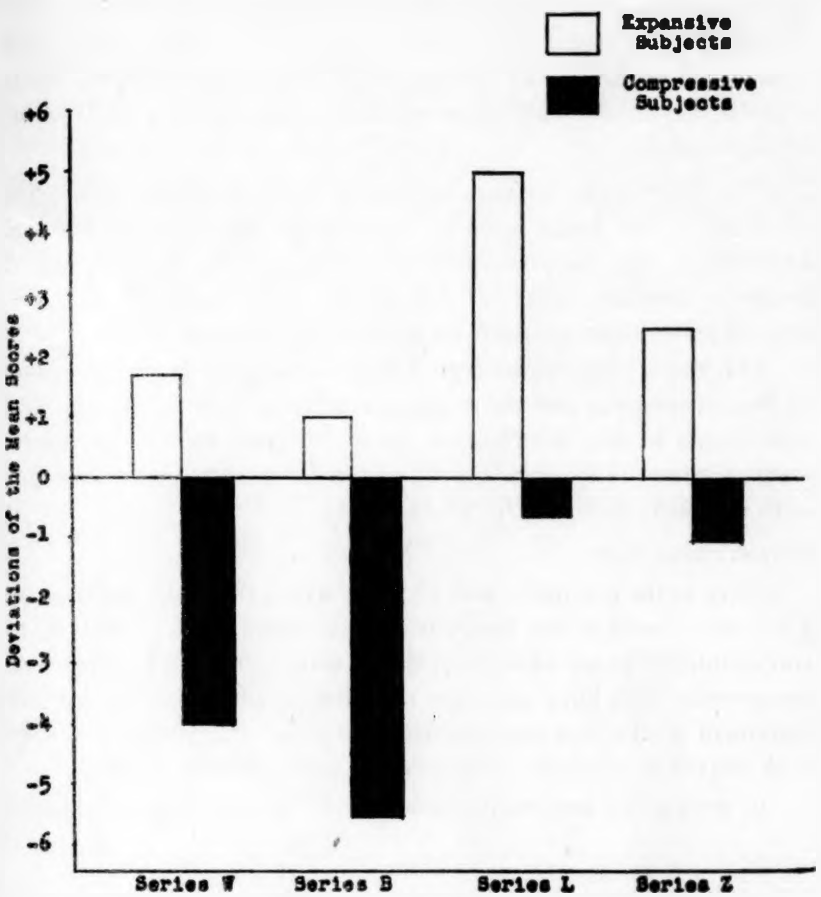


FIG. 2. Deviations of mean scores for the expansive and compressive group in each clairvoyance series. (The deviations are found by subtracting the mean chance expectation, 40.00, from the mean score per set of drawings.)

the compressive group is lower than chance expectation. This is even more clearly indicated in the graphical presentation of these results in Figure 2.

(A mean score above chance indicates the presence of similarities between stimuli and responses, while a mean score below chance is indicative of dissimilarities.)

When the ESP scores of the expansive subjects alone are con-

sidered, we find that the difference between their mean score of 41.88 and the chance expectation of 40.00 gives a t of 1.64. This t has a tabulated probability of .05, but since it has to be corrected, it has a probability of .10 which does not achieve the criterion of 1% level of significance.

The ESP scores for the compressive subjects give a mean score of 37.45. This mean score is significantly different from chance expectation since the probability of getting a t this large (2.71) or larger in absolute value (2P) is .0078. This indicates an avoidance of relationship between the stimuli and the responses.

The statistically significant difference between the mean scores of the compressive and the expansive subjects ($t = 3.00$) has thus been found in data which were, in the original report, not significant in terms of the total results of the direct matchings of stimuli with responses intended for them.

Displacement Data

Since in the original report of these series, the main evidence for ESP was found in the backward displacement data, it may be of special interest to see what effect the division according to expansion-compression will have upon the distribution of the scores for displacement to the just-previous stimulus (that is, Response 2 rated with respect to Stimulus 1, Response 3 with Stimulus 2, etc.).

In testing for backward displacement, the first response of each

Table 2

TOTAL SCORES FOR DISPLACEMENT BACKWARD AS DIVIDED BY THE EXPANSION-COMPRESSION CRITERION

Series	EXPANSIVE SUBJECTS			COMPRESSIVE SUBJECTS			Difference Between Mean Scores of the Two Groups	2P
	N*	Mean Score on ESP Test	t	N*	Mean Score on ESP Test	t	t_d	
W.....	15	28.53	.75	9	33.56	1.89	1.73	.096
B.....	10	30.30	.16	14	37.43	5.59	3.26	.0037
Z.....	14	30.07	.05	10	34.20	3.21	2.10	.047
Total...	39	29.54	.46	33	35.39	6.07	4.37	.0001

*N refers to the number of subjects each of whom made one set of four drawings.

subject has to be omitted since there is no "just-previous" stimulus with which to compare it; therefore the mean chance expectation per subject in this case will be 30.00.

In Table 2 it can be seen that the compressive group gave highly significant evidence of displacement with a t of 6.07. On the other hand, the expansive group scored virtually at chance ($t = .46$). The t of the difference between the mean scores of the expansive and

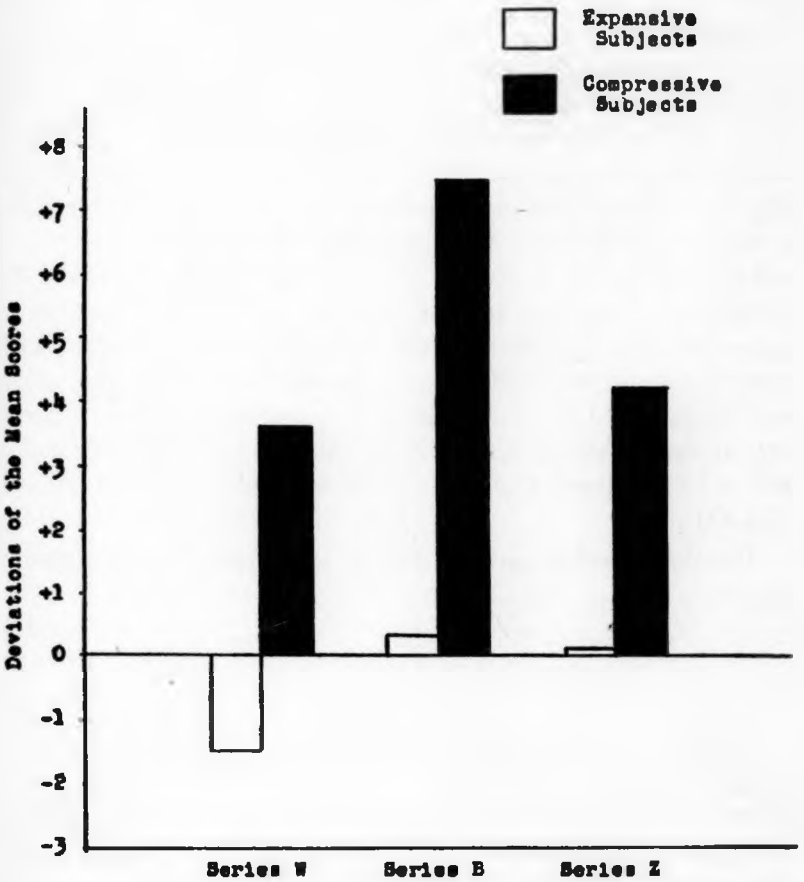


FIG. 3. Deviations of mean scores for the displacement data of the expansive and compressive subjects in the clairvoyance series. (The deviations are found by subtracting the mean chance expectation, 30.00, from the mean score per set of drawings.)

compressive subjects is quite significant, being 4.37, and having a probability (2P) of .0001 that such a difference could have occurred because of fluctuations of random sampling. That the individual series show similar trends may be seen from the table and from the graphs in Figure 3. (It was explained in the original report on these series that it is not possible to obtain displacement scores for Series L since the experimenter did not keep a record of the *order* in which the responses were made.)

Combined Results

It is recognized of course that the scores for backward displacement are not independent of the direct matching scores already cited. If, however, we desire to note the total effect of the expansion-compression criterion as a discriminative measure, we can sidestep the issue of the interdependence of direct and displacement scores by hypothesizing that expansive subjects will score directly and that compressive subjects will give evidence of backward displacement. This can be tested by combining the direct scores of the expansive subjects with the backward displacement scores of the compressive subjects (taking into consideration the proper adjustment necessitated by the dropping of one response from each subject in the backward displacement data). The total mean score obtained thus gives a t of 5.44, which has a probability of less than .000,000,1.

For the individual series, a similar treatment gives the following figures:

Series W: $t = 2.26$	$P = .017$
Series B: $t = 3.41$	$P = .0013$
Series Z: $t = 2.65$	$P = .0073$

Thus we see that the expansion-compression criterion very effectively focuses the evidence for ESP in these clairvoyance series: by this measure two series are independently significant, while the third is suggestive with a probability of less than 1 in 50. (Comparable data are not available for Series L.)

The Judgment of Expansion-Compression

In judging expansion-compression, subjective elements will ob-

viously play some part. Elkisch mentions that there are great differences in the degree to which a certain quality may appear and that such differences may be given different emphasis by various judges. She points out that there is little possibility of evading such subjective elements; and, moreover, she warns against trying to apply the criterion in a mechanical or wooden fashion.

A similar difficulty has been found in other types of psychological research. For example, in studies of the judgment of personality from facial expressions, it has been found that people differ greatly in their ability to make correct or valid judgments. In some of these researches naïve laymen or persons having an active interest in painting made the best judgments (8); whereas, scientists and psychologists in particular made the poorest judgments (8, 20). It was suggested that the latter result might be due to the fact that "the mental activities are constrained by calculating hypothesis and inferential reasoning" (20).

In spite of the great differences in ability to judge various phases of personality, it has been found possible to train judges to a more accurate perception and more valid judgments (1, 2). Such training brings about more precise understanding of terms used, calls attention to the most important aspects of the features to be judged, and gives the proper perspective to details by means of concrete examples.

Because the judgment of expansion-compression is such a delicate appraisal of quality and because there are such individual differences in ability to judge such qualities, complete agreement among judges is an almost impossible attainment. And yet it is obviously necessary to obtain some idea of the reliability of this criterion if it is to be of use generally and if the hypothesis of ESP on the part of the judge is to be ruled out.

The best available evidence on this point of agreement among judges has been discovered too recently for other than brief mention here. A recent attempt to repeat these experiments (to be reported in a later issue of the JOURNAL) has yielded a much higher correlation between judges of expansion-compression. In this case Dr. Stuart and I made the judgments on 32 sets of drawings and our judgments were found to give a correlation coefficient of +.66,

which is highly significant. This indicates that it is entirely possible to reach a degree of agreement which is fairly high for this type of judgment.

There is, however, a phase of the present experiment which also offers evidence on the point of agreement among judges, although it is not so conclusive. Mrs. Dorothy H. Pope of the Parapsychology Laboratory was asked to undertake the task of judging expansion-compression in the 384 responses of these four clairvoyance series. Although Mrs. Pope had never had any experience in a task of this specific sort, she had shown appreciation of subtle relationship in judging similarities between responses and stimuli in other experiments. The relevant sections of the Elkisch monograph were given her to read; and, in addition, we went over the material verbally together and discussed the illustrations in the Elkisch report. After this study, Mrs. Pope evaluated the drawings of Series W. Her judgments were then compared with mine and the sets upon which we did not agree were discussed. (None of her judgments were changed, of course.) After this additional bit of instruction, she judged the form quality in the drawings of the remaining three series.

The ESP scores of the two groups of subjects—expansive and compressive—as separated by Mrs. Pope are:

Direct Matching Data		Displacement Data	
Expansive Group:	40.26	Expansive Group:	41.63
Compressive Group:	38.60	Compressive Group:	42.78

These mean scores differ in the same direction as those given in Tables 1 and 2: in the direct matching data the score of the expansive group is higher than that of the compressive group, while in the displacement data the reverse is true. The differences between these mean scores, however, are not at all as large as those given by my judgments of the data. Thus there are obviously differences of opinion between the judges. In fact, a Pearson product-moment correlation of the expansion-compression scores assigned to each subject by the two judges gives a coefficient of only +.40, which is a statistically significant correlation coefficient but one which is indicative of low reliability of the measure.

A better test of the discriminative value of the expansion-compression criterion may be found, however, by considering only those cases upon which the two judges agree. By this means we should expect the most clear-cut cases to be found if the expansion-compression criterion is actually descriptive of recognizable qualities. Therefore, the ESP scores were compiled for those subjects upon which the two judges agreed as to the presence of expansion or compression in their drawings. The mean scores for the two groups are presented in Table 3.

Table 3

COMPARISON OF THE MEAN SCORES OF THE TWO GROUPS OBTAINED WHEN TWO JUDGES AGREE AND THOSE OBTAINED FROM ONE JUDGE ONLY

A. Direct Matching Data

Conditions	Expansive Subjects		Compressive Subjects		Difference Between Mean Scores
	*N**	Mean Score	*N**	Mean Score	
Mrs. Pope and B.M.H. agree	23	43.17	35	37.77	5.40
B.M.H. alone	41	41.88	55	37.45	4.43
Remainder of B.M.H.'s results upon which Mrs. Pope did not agree	18	40.22	20	36.90	3.32

B. Displacement Data

Conditions	Expansive Subjects		Compressive Subjects		Difference Between Mean Scores
	*N**	Mean Score	*N**	Mean Score	
Mrs. Pope and B.M.H. agree	21	28.76	19	35.00	6.24
B.M.H. alone	39	29.54	33	35.39	5.85
Remainder of B.M.H.'s results upon which Mrs. Pope did not agree	18	30.44	14	35.93	5.49

*N refers to the number of sets of four drawings each.

There it may be seen at once that the differences between the two groups are greatest when the two judges agree upon the quality expressed. This is the case for both the direct matching data and

the displacement data. In each case the difference between means when the two judges agree is greater than those of either judge alone. In spite of the reduced number of cases, both of these differences are statistically significant. A t of 2.83 was obtained for the direct data, while a t of 3.03 was found for the displacement data.

In order to show that this effect is not a function of the sampling of a batch of data known to be significant, additional information is given in Table 3. If we consider my division of the cases upon which Mrs. Pope and I did not agree, we find a difference in the predicted direction, but one which is not nearly as large as that obtained when we both agree. This is true for both the direct and displacement data. Therefore it seems safe to say that the statistically significant difference obtained when both judges agree was not produced by sampling data any division of which might have produced a significant difference.

The fact that Mrs. Pope's judgments of expansion-compression separate the data in a manner similar to that in which mine divided them shows that the categories as defined are meaningful and the qualities recognizable. When the two of us agreed, the difference between the two groups of subjects was even more striking. In view of other psychological studies, it is quite probable that Mrs. Pope's judgments would have been even more discriminative if she had been given more training. It would seem that this training would produce the greatest advantage if it were continued to the point where the judgmental process becomes almost automatic or instantaneous.

Discussion of Counterhypotheses

That the conditions surrounding the acquisition of the experimental observations were entirely adequate for the preclusion of sensory cues may be seen from the original report of the series involved (18). But in brief, it may be recalled that the stimuli were sealed within opaque envelopes, and that the matching of the stimuli and responses was also done under carefully guarded conditions so that there were no cues by means of which the judges could have inferred the true correspondences.

All checking and statistics pertaining to these series have been independently carried out by at least two persons. Thus the results could not be due to errors in computation. Likewise all statistics pertaining to the division of ESP scores by the expansion-compression criterion have been independently checked. And, as was explained earlier, the results of the expansion-compression analysis were determined by Dr. Stuart and myself in a two-experimenter procedure with shared responsibility.

Before accepting the proposition that the mean scores are fairly indicative of the degree of relationship between stimuli and responses, we must consider certain counterexplanations. We must also consider counterexplanations which involve the ESP hypothesis. One such hypothesis has to do with the fact that the judges' matchings may be not only indications of objective relations between stimuli and responses, but may also involve ESP ability of the judge in perceiving relations. The judges may conceivably use ESP to match a given response with the stimulus for which it was intended. This possibility has been discussed at length in the original reports, and there are several arguments against it. One serious complication for the "ESP-in-the-judge" hypothesis lies in the significant displacement data. To account for these the judge would have to know by ESP the true order of both stimuli and responses and then *shift* his matchings one place so that each response would be matched best with the just-previous stimulus. It is difficult to understand what motivation would prompt a judge to do this. Moreover, the scores used are made up of the judgments of *two* judges. Thus, to get the consistent results found here, both judges would have had to show backward displacement. This is further complicated by the fact that the same two judges did not judge any two series.

The best argument against the hypothesis that the judges used ESP is provided by the results of the present study. Here a judgment has been made of subject performance according to the expansion-compression criterion. For our significant differences to be due to judges' ESP ability, we would have to construct a very complicated hypothesis in which the judges of the ESP scores *foresaw* the eventual application of the expansive-compressive criterion and again

used ESP in their matchings so as to give one group high scores and the other low scores.

Acceptance of such an hypothesis would mean imputing, not to one judge but to several, a latitude of ESP abilities not yet proved experimentally. Thus we may conclude that there is no reasonable likelihood that the results of the present study are due to ESP judgments on the part of the judges.

Extrasensory perception might be supposed, however, to have been an important factor in my judgments of form quality. One might assume, for example, that I had some extrasensorial knowledge of the ESP scores of each subject and on the basis of this information unconsciously separated the subjects into two groups one of which gave a positive score while the other was negative. In addition to attributing to me extrasensory powers which have never shown up in other tests, this hypothesis is not borne out by the fact that Mrs. Pope's judgments of form were similar to mine. Now one would have to postulate that both Mrs. Pope and I used ESP to separate the subjects.

The more recent experiment referred to earlier in which the judgments of Dr. Stuart and myself showed a higher correlation affords additional evidence that the criterion of expansion-compression does describe such recognizable qualities that it is unnecessary to have recourse to the weak hypothesis that two judges unconsciously used ESP in making the form judgments.

The significant differences obtained could not be due to any spurious factors in the application of the expansion-compression criterion. As was stated earlier, at the time I analyzed the responses for formal qualities, I did not know the ESP scores of any subject.

There is another type of counterexplanation which would suppose that it was the *targets* which were expansive or compressive, and not the subjects. For example, is it not possible that some subjects made expansive drawings, not because of any expansive quality in their own approach to the test but because of their ESP ability? Other subjects not having ESP ability may or may not have made expansive drawings. One might expect that by chance half would make expansive drawings and half would make compressive drawings. According to such an hypothesis, the expansive and com-

pressive groups of subjects differ only by virtue of the fact that one group had ESP of targets which happened to be expansive, while the other had no ESP ability. This division, therefore, would be nothing more than a separation of the good and poor ESP subjects and would be in no way tied up with a projection of the subjects' own personalities.

This complicated hypothesis can be met on a number of grounds,[3] but it can most easily be disposed of by saying that an analysis of the form quality in the stimuli was made to test this hypothesis. Three judges rated all the stimuli according to the expansion-compression criterion. An analysis of the ESP scores for the expansive and compressive groups of subjects for the "expansive" and "compressive" stimuli showed that the hypothesis did not hold. Expansive subjects hit "expansive" and "compressive" targets and compressive subjects hit both types of targets also. The quality judged to be present in the stimuli did not seem to be bound up with the type of group which obtained the best scores on that particular kind of picture.

Thus the only reasonable explanation of these results is that an analysis of expansion-compression in drawings does effectively separate subjects of ESP tests into two groups with significantly different scoring levels.

Discussion of Findings

The expansion-compression criterion has proved to be a measure capable of separating high-scoring subjects from low-scoring subjects in ESP tests with enclosed drawings. The subjects rated as expansive gave scores above chance, while the compressive ones scored below chance expectation on the ESP test. The statistically significant difference of 3.00 between the two groups of subjects can only be attributed to differences in ESP performance within the series. Moreover, this significant difference has been found in series which were originally believed to obtain no evidence of ESP in terms of the direct matching scores.

The division of the backward displacement scores according to

[3] For full discussion of the evidence against this hypothesis together with the figures of the analysis discussed, the original thesis deposited in the Duke University Library should be consulted.

expansion-compression shows an even more striking difference (t = 4.37), with the compressive group alone contributing highly significant evidence (t = 6.07) of backward displacement.

The criterion, then, has proved valuable as an empirical finding. But besides its having value as an empirical finding, it would appear that the discovery of the *meaning* of this measure in terms of subject performance would be a far more important addition to our knowledge. It may be years yet before the entire picture concerning the interpretation of this discovery is completed by future researches, but it may be well to try to draw into focus the areas of certainty and uncertainty.

The expansion-compression criterion was originally utilized to discriminate between adjusted and maladjusted children. A similar study with adults by Schmidl Waehner (19) has shown a significant relationship between a diagnosis of maladjustment on the basis of free drawings and a diagnosis of maladjustment based on the Rorschach test. The studies of Allport and Vernon (3) and those of Werner Wolff (20) indicate that expressive movements are consistent and that they do reveal personality traits of the subject—they are "representative of the inner personality of the subject."

We must remember, however, that there is an important difference between our ESP series and the conditions in the experiments on expressive movements. In the psychological studies the movement diagnosed was "spontaneous." An attempt was made to record those expressions which are "dependent less upon external and temporary conditions than upon enduring qualities of personality" (3, p. 23).

We do not know enough yet about how the situation surrounding an ESP experiment affects the personality. One might suppose that an individual will react similarly in a variety of situations in regard to expressive movements. And yet the ESP test situation is a very unusual situation to most subjects who have never been tested before. It is quite conceivable that the "normal, adjusted" person reacts in a manner which is quite contrary to his natural manner of expression. As a matter of fact there is no "normal" way to react to the ESP test. Some may find the situation gamelike and amusing. Others may evidence various degrees of frustra-

tion. A few show signs of anxiety. Expansiveness of response might be normal for a person enjoying a game, and compressiveness of response, normal for a person "taking a Psychological Test." Thus it seems possible that drawings obtained in an ESP test may be more dependent upon temporary conditions "than upon enduring qualities of personality." Until more information is available on this point, it is certainly not warranted to conclude that compressive subjects in these series are generally maladjusted.

In attempting to find an explanation for the results of this study, we must include in our thinking the displacement results. This raises a problem in itself since we have no adequate explanation for the occurrence of displacement in these series. Backward displacement can hardly be conceived to be a necessary accompaniment of clairvoyance drawings tests. Backward displacement has been found by Carington in his GESP drawings experiments (6), but there are no clairvoyant drawing experiments available for comparison. In card-calling tests of the clairvoyant type there was no displacement found in the three series analyzed by Russell (14).

Although at this stage we cannot be sure that the ESP process operating in drawings tests follows the same trends as it does in card-calling tests, this at least seems a likely supposition. Thus until further research is conducted, it seems reasonable to consider that the displacement phenomenon in the present clairvoyance series is probably due to some peculiar circumstances connected with the four series. For example, Stuart (18) has suggested that the just-previous stimulus may become salient by being in the hands of the experimenter who is the dominant figure in the experiment or that it becomes so by virtue of some latency effect. We will have to wait upon future experiments for the final answer to this problem.

In trying to reconstruct what occurred in the experimental sessions, we must keep in mind that ESP is essentially an *unconscious* process. Although it can be consciously directed to some degree, it seems to be influenced by unconscious motivation and shows up in ways unintended by the subject. Thus in the present series, one might suppose that something like this took place: some of the subjects did not feel particularly disturbed or constrained about this "impossible" drawing task which they had been asked to do, and

they drew freely and expansively, and actually achieved a fair degree of success in hitting the assigned targets. On the other hand, other subjects felt frustrated or inhibited by this task which they perhaps took too seriously. They drew very timidly and with a great deal of uncertainty and perhaps even anxiety. With something akin to relief, they returned the first envelope and its enclosed stimulus to the experimenter and then began on the next task. Now as they took up this new envelope they experienced again a feeling of inhibition and so again they drew very tense, cramped, tiny figures. But by one of those tricks of the mind, their drawings significantly resembled the previous stimulus. It may be that something happened that is analogous to name-forgetting in which the harder one tries and tries to bring the desired name to mind, the more the elusive word seems to slip from memory; yet the moment one turns his attention to something else, the sought-for word comes spontaneously to mind. So, after a subject has finished with one envelope upon which he has been working with some anxiety and uncertainty, the sudden reduction of these feelings with respect to the just-previous stimulus may allow for the salient features in that picture to rise spontaneously in the subject's mind as he is struggling with the next "impossible" task, and so on through the remainder of the test.

Such lines of thought are speculative and are useful mainly as they serve to guide experimentation. Further discussion aimed at explaining these results must wait upon the reporting of the analysis of the GESP series, since an adequate explanation of results must include the results of both types of ESP test procedure.

The discovery of one form quality which discriminates between high- and low-scoring ESP subjects has initiated a number of follow-up researches which will be reported in time. Some of these, together with other suggestions for possible lines of further research, will be discussed in the report of the GESP series.

REFERENCES

1. ALLPORT, F. H. *Social Psychology.* Boston: Houghton Mifflin, 1924.
2. ALLPORT, G. W. *Personality, a Psychological Interpretation.* New York: Henry Holt, 1937.
3. ALLPORT, G. W., and VERNON, P. E. *Studies in Expressive Movement.* New York: Macmillan, 1933.
4. ANASTASI, A., and FOLEY, J. P., JR. A survey of the literature on artistic behavior in the abnormal: III. Spontaneous productions. *Psychol. Monogr.*, 1940, **52**, 1-71.
5. ———. A survey of the literature on artistic behavior in the abnormal: IV. Experimental investigations. *J. gen. Psychol.*, 1941, **25**, 187-237.
6. CARINGTON, W. Experiments on the paranormal cognition of drawings. *J. Parapsychol.*, 1940, **4**, 1-129.
7. ELKISCH, P. Children's drawings in a projective technique. *Psychol. Monogr.*, 1945, **58**, 1-31.
8. ESTES, S. G. Judging personality from expressive behavior. *J. abnorm. soc. Psychol.*, 1938, **33**, 217-36.
9. HUMPHREY, B. M. ESP and intelligence. *J. Parapsychol.*, 1945, **9**, 7-16.
10. ———. An exploratory correlation study of personality measures and ESP scores. *J. Parapsychol.*, 1945, **9**, 116-23.
11. MONROE, R., LEWINSON, T. S., and WAEHNER, T. SCHMIDL. A comparison of three projective methods. *Character & Pers.*, 1944, **13**, 1-21.
12. PEARSON, K. (Ed.). *Tables for Statisticians and Biometricians.* Part I. (3rd ed.) London: Cambridge University Press, 1930.
13. RIDER, P. R. *An Introduction to Modern Statistical Methods.* New York: John Wiley, 1939.
14. RUSSELL, W. J. Examination of ESP records for displacement effects. *J. Parapsychol.*, 1943, **7**, 104-17.
15. SARGENT, H. Projective methods: their origins, theory, and application in personality research. *Psychol. Bull.*, 1945, **42**, 257-93.
16. STUART, C. E. An ESP test with drawings. *J. Parapsychol.*, 1942, **6**, 20-43.

17. ———. A classroom ESP experiment with the free response method. *J. Parapsychol.*, 1945, **9**, 92-105.

18. ———. An ESP experiment with enclosed drawings. *J. Parapsychol.*, 1945, **9**, 278-95.

19. WAEHNER, T. SCHMIDL. Formal criteria for the analysis of children's drawings. *Amer. J. Orthopsychiat.*, 1942, **12**, 95-104.

20. WOLFF, W. *The Expression of Personality.* New York: Harper, 1943.

Parapsychology Laboratory
Duke University
Durham, North Carolina

ON THE NATURE OF PSI PHENOMENA

By R. H. Thouless and B. P. Wiesner[1]

It is assumed in this paper that the reality of the facts commonly referred to as those of "paranormal cognition," "extrasensory perception," or the "psi phenomena" have now been established by experimental research beyond any possibility of reasonable doubt (4, 5, 6, 7). We also consider that the reality of psychokinesis is equally well established (3). Furthermore, it seems to us that these researches have established certain characteristics of these phenomena (to be mentioned later) which provide a basis for some amount of tentative theoretical speculation.

The standpoint assumed is that already sketched by one of us (9). It has been proved that reactions can be obtained (as in card-guessing experiments) which are determined by facts not known to the subject by any normal means, neither by sensory perception nor by rational inference. In some cases, knowledge of the fact by some other person may be a condition favorable to the appearance of the reaction, but it does not appear to be a necessary condition. This proposition would be generally accepted by workers in psychical research although there is some controversy about it. There is no doubt that distances of even hundreds of miles between the event and the consequent reaction may not prevent the appearance of a successful reaction, and there is no evidence that such great distances are even an unfavorable condition (1, 6). It also appears that a reaction may be determined by a future fact which also may not be known to the subject by any normal means. It seems preferable, as a provisional hypothesis, to regard these different phenomena not as effects of different paranormal capacities (telepathy, clairvoyance, and precognition) but rather as manifestations of one single paranormal capacity whose characteristic is that reactions may be

[1] Dr. R. H. Thouless of Cambridge University is already well known to our readers through his articles. This is the first contribution to the Journal by Dr. B. P. Wiesner, medical specialist and research worker, who is now collaborating with Dr. Thouless on research in parapsychology.—Ed.

determined by events which cannot influence the subject in any normal way and which need not show the normal spatio-temporal relations between event and reaction; that is, they are not necessarily limited by the normal condition that a reaction can only follow from an event if some physical consequence of that event is at the same place as the subject at some time preceding the occurrence of the reaction. For this paranormal capacity we use the term "psi," and we refer to the observed facts of psi as the "psi phenomena."

We are, no doubt, still far from the ideal position in which research will, as we hope, give us a few principles from which all facts of the psi phenomena can be deductively established. It seems, however, that already we have reached a position in which we can begin to form postulates from which consequences may be deduced and that these may serve as a guide for further experimentation. In forming such postulates, we must begin by making them as general as possible and with the least possible number of specific assumptions.

Any theorizing about these phenomena must start from the fact that psi-determined behavior, like perceptually controlled behavior, is, at least in some degree, correlated with events in other parts of the universe. We may express this fact of correlation by saying that some part of the universe is projected by some means on the behavioral system of the organism. The conscious process of perception is a familiar means of such projection. The kind of projection that produces sensory perception might have turned out to be the only kind of projection, but the discovery of the psi phenomena show that this is not the case. Even when there is no resultant conscious process we must postulate a projection of part of the external universe on to the behavioral system of the organism when (as in psi-determined responses) there is more than chance correspondence between the behavior of the organism and some part of its environment.

Let us consider two possibilities as to the nature of the projection of segments of the universe upon the behavioral system of an organism: such projection might be wholly or not at all conditioned by the spatio-temporal contiguity of the projected segment to the organism in question. These two possibilities may be regarded as polar opposites. It is not necessary to suppose that any particular

projection must be either completely limited to a spatio-temporally contiguous segment of the universe or completely free from such limitation. It may lie anywhere between these extremes.

The first of these opposite possibilities is that familiar to us in sensory perception. Here the projected part of the universe is one determined by its position in space and time relative to the organism in a manner dependent on the nature of the sensory mechanisms and on the physical properties of the stimuli to which these respond. Projection of segments of the universe limited by contiguity with the organism in space and time might also be brought about by other means than the sensory mechanisms; projection of the psi type is not found necessarily to involve complete dislocation in space and time.

The second possible type of projection is that in which there is no such rigid limitation of the projection to that segment of the universe which is in spatio-temporal contiguity with the organism. This form of projection could take place only by means of some other mechanism than that of the sensory apparatus. There are obvious difficulties (to be considered later) in the way of accepting the possibility of projection of this type, but acceptance is made necessary by well-attested experimental facts of psi phenomena (**1, 4, 7**).

We may consider the extreme form of this type of projection as one in which there is correspondence between the behavior of the organism and events not restricted to any particular part of the universe. The event projected might be one happening anywhere in the universe at any time irrespective of the location in space and time of the organism.

It is true that if the behavioral system of an organism were influenced by random projections from all parts of the universe (whether remote or near in time and space), it would follow that such an organism would not be related in any serviceable way with its environment. This, however, does not imply that such a system of projection is impossible, but only that the organism would have the property of being unable to make adaptive responses to environmental demands. Adaptive responses to environmental demands become possible only when there is some degree of restriction of

projections, or at least of the organism's responses to them, to spatio-temporally contiguous parts of the universe.

The observed facts of the psi phenomena do not, however, require us to believe in the occurrence of this extreme form of totally unrestricted projection, but only in the occurrence of a form of projection considerably less restricted in space and time than are the projections of sensory perception. If we suppose that the first stages of life may have consisted of primitive organisms receiving projections at random from the rest of the universe, we must also suppose that evolutionary development had, as one of its accompaniments, the progressive restriction of projections to those from more or less spatio-temporally contiguous events which alone can lead to serviceable responses.

Another objection which may be urged against the possibility of even the relatively unrestricted type of projection which the facts of psi phenomena compel us to accept is that we cannot imagine any mechanism by means of which such projection could take place. This certainly is true, but it may only be because, when we try to imagine a mechanism of psi projection, we find ourselves unable to imagine it otherwise than as working in the same sort of way as sensory perception—by radiations from the object falling on a sense organ, etc. It seems clear that the mechanism of psi projection is not of this kind. At some stage of our knowledge of this subject, we must hope to be able to solve the problem of the mechanism of psi projection, since it is evident that no projection can take place unless there is some means by which it takes place. At present, however, we are not able even to guess in what direction or in what terms a solution could be found. We must therefore be content to consider possible relations between an organism and the rest of the universe without concerning ourselves at this stage with the mechanism of these relations.

If projection can take place by means other than the sensory apparatus, it is possible that projection by means of the senses is superimposed upon a more or less developed nonsensory system of projection. The behavior of the organism may thus be determined in part by projections of segments of the universe beyond the reach of the senses and in part by projection of the sensory-perceptual

type. Any contiguous segment might therefore be projected both by psi and by sensory means.

It is interesting to consider what may be the functional and evolutionary significance of the relationship between these two different forms of projection. It is obvious that any useful adaptation of the behavior of the organism to the conditions of its environment requires some degree of limitation of its behavior responses to those determined by events in the contiguous environment. This is usually, although perhaps not entirely, achieved by the limitations inherent in sensory perception and by the exclusion of psi projections that come from noncontiguous parts of the universe. The sensory mechanisms may be regarded as having the function not only of acting as a means of projection of the contiguous environment but also of giving preference as a determinant of behavior to this projection rather than to those of the psi function which tend to remain relatively unrestricted and random. In other words, the role of sensory perception may be considered to be that of replacing a more primitive (and generally less biologically useful) kind of projection and of more or less completely suppressing it (9). It is not essential for this purpose that sensory perception should have the capacity for accurate localization of events so long as some order is brought into them. It may, for example, create order by introducing a time sequence although this order may not correspond exactly to the physical time relationships of the external events projected. It is well known that the power of lower organisms to locate environmental events in space is very limited, and this may be true of time.

There is no reason for supposing that the development of sensory perception necessarily abolishes the more primitive psi function; and, as has already been mentioned, there is much empirical evidence that it does not. All that is implied by the hypothesis which is being discussed here is that the function of the sensory apparatus is to decrease in the course of evolution the occurrence of behavior responses to more or less random projections of the universe which result from the primitive psi function and gradually to replace these by responses to that limited segment of the universe which is in spatio-temporal contiguity with the organism. We can imagine that

evolution might have proceeded along different lines by producing a limiting mechanism for psi so that in psi projection there would be adequate discrimination between contiguous and noncontiguous parts of the universe. The fact that some "sensitives" can use psi in a practically useful manner indicates that such limitation is possible (2).

We have used the term "psi" for the primitive mode of extrasensory projection of the universe on the organism. We may also use the term "S" for that mode of projection which is the result of the activity of the sensory apparatus. As evolution proceeds we may suppose that a stage is reached at which the psi function is, for all ordinary behavior of the organism, replaced by the S function. This does not imply the abolition of psi but the more or less complete suppression of its effects on behavior in normal circumstances. In the highly developed behavioral system of man, S may appear to have succeeded in continuously and firmly relating the organism to a contiguous segment of the universe and in preventing, in all ordinary circumstances, the intrusion of projections of noncontiguous segments into his behavioral universe.

The hypothesis that the psi type of projection is being gradually suppressed in the course of evolution by the S type may be tested in two ways. First, we may try to discover whether it is compatible with known facts; secondly, we may try to deduce from it expectations concerning phenomena which are not yet known.

Let us consider the conditions under which the psi function might be demonstrated by card-guessing experiments of the type originated by Rhine. If the subject's guesses were wholly determined by psi projections of the hypothetical primitive, unrestricted type, these experiments would yield no significant results, for a set of guesses which were equally likely to be determined by the order of distribution of all packs of test cards anywhere in space and time would show only the same sort of mean chance correspondence with the particular pack used as standard as would a set of guesses obtained without any operation of psi. In other words, complete lack of limitation of psi would lead in such experiments to the same result as its complete absence. If, however, the psi-determined responses are sufficiently limited to that pack of cards which is used as standard, then successes deviating from mean chance expectation

will occur. This is what has been found to happen in many experiments (1, 5, 6, 7, 8). It is apparent, then, that this limitation of psi is necessary for its demonstrability and that subjects may fail to show significant results, not because there is no psi projection, but because the psi function is insufficiently limited by the suppression of responses to projections from remote sources.

It is not necessary to suppose that the suppression which determines the relative effectiveness of the remote and of the contiguous on the behavior resulting from psi projections takes place always in the same way for different subjects or for the same subject at different times. For any subject, it may be that the projection of any segment of the universe occurs only for a short period. Even where psi is limited in any particular way during a prolonged period (as must be the case in a successful card-guessing experiment) the limitation might persist with respect to location but not with respect to time. Thus a pack of cards may be projected by psi for a period of some minutes but the projection of each particular card may oscillate on either side of the present moment so that the actual card projected is sometimes a card somewhat earlier in the pack than the one turned up and sometimes somewhat later. If, as in the Duke University experiments, subjects are regarded as successful only if they guess correctly the card exposed, it may well be that there are other subjects in whom psi is also effective in determining card guesses. The subjects who guess correctly the card exposed are presumably those in whom the relatively remote psi projections are most effectively suppressed. In others, psi may be less easily demonstrable because their psi-determined guesses are continually varying with respect to time. They may, for example, now receive a projection of a card three places ahead and immediately afterwards a projection of one three places behind. The methods for the detection of psi which have been generally used have been concerned, if this hypothesis is correct, with discovering manifestations of a particular type of limitation of the psi function. Some method of scoring hits on cards turned up over a period of time (for example, regional scoring) should reveal the presence of the type of psi projection which results from incomplete suppression.

The hypothesis which is here sketched out suggests certain lines

of experiment. It suggests, for example, that we should seek to discover what conditions as to the functioning of S will produce optimal psi results. If sensory perception is a later acquired process normally suppressing the effects of psi, it is to be expected that the presence of any factor in the experiment favoring S would be detrimental to psi. It is, for example, generally supposed that an experimental subject will score better on card-guessing experiments if there are any sensory cues which could help him to get the card right by sensory perceptual methods. The above line of thought suggests, however, that the presence of such sensory cues might be detrimental to psi activity and that it might make a successful psi subject less successful.

It also suggests that information given to the subject as to his successes may be detrimental to psi success and not (as seems generally supposed) favorable to it. If a subject is learning, let us say, to throw darts at a target, the condition of improvement is that he shall know when he has failed and when he has succeeded in his object. This is a type of behavior conditioned by the projections of the sensory perceptual apparatus. Information as to success, if favorable to S responses, may be expected to be unfavorable to psi responses. It is therefore to be considered whether optimal conditions for psi experiments cannot best be obtained by the opposite method of giving the subject no information as to success or failure.

If this is so, it may be the explanation of the common observation that subjects showing successful results in card-guessing experiments get progressively worse as a series of experiments goes on and finally score nothing beyond mean chance expectation. The effect of the course of experiments seems to have been not to train the psi reactions of the subjects but to train their suppression of psi reactions.

A further point of interest bearing on our hypothesis is the observation that, with some subjects, when the attempt is made to guess the card turned up by the experimenter, successes are obtained when judged by some other card in a fixed relation to the one turned up (the next card or the next but one) (**1, 4, 7, 8**). In other words, the subject fails in the task he is attempting to perform but seems to be succeeding in some other task which is not his object. This may be explained by supposing that the suppression of psi is effective

in the task towards which the effort of the subject is consciously directed but that psi projections may escape this suppression if they refer to some other object than that at which the conscious aim of the subject is directed. It is as if there were a censorship directed towards the prevention of any influence on behavior by psi projections of the part of the environment with which the conscious behavior of the subject is at that moment concerned but which is relatively indifferent to the intrusion of psi projections referring to some other part of the environment.

For much the same reason, it is to be expected that psi projection may most easily influence those elements of the subject's behavior which are not those towards which his conscious effort is directed. Thus psi projection effectively censored from influencing a subject's conscious behavior might be found to influence involuntary movements which are not cognized by him.

These considerations suggest two principles which may be employed in psi experimentation, either separately or together. The first is that the conscious activity of the subject should be directed towards some other task than that in which the experimenter intends to look for the effects of psi. The second is that the experimenter should look for the effects of psi in some other elements of the subject's behavior (such as involuntary movements) than those towards which his conscious effort is directed.

That these principles are not merely fanciful is strongly suggested by the observation that from the remote past both principles have been employed by prophets and soothsayers who have made their living by utilizing psi projections. Examples of the conscious efforts of the prophets being directed towards other tasks than that on which psi projections are desired are to be found in the examination of the entrails of animals or the arrangement of tea leaves in a saucer or in the lines in the palm of a hand. The explanation which our hypothesis suggests of the effectiveness of such methods is that the prophet examining the entrails of an animal will be effectively prevented, by the automatic suppression of psi projections, from obtaining knowledge by psi of any information about those entrails. This, however, he does not desire; what he wants is a psi projection of the future course of a battle. If he consciously directed his

efforts towards foreseeing this, the automatic suppression of such psi projections would come into effect. If, however, his conscious efforts are directed towards examining entrails, there is no reason why psi projections of the future course of the battle should be suppressed. He has, as it were, hoodwinked the censorship.

Examples of the second principle (psi affecting involuntary movements) are also to be found in traditional practice. In table-turning and the use of the Ouija board, involuntary movements are used as a means of conveying information derived from psi projections. It is true that in table-turning, at any rate, psychokinetic movements may also play a part. If this is the case, it may be regarded also as an example of the operation of the same principle that psi can most easily affect movements other than those consciously controlled. When the subject is in a state of trance or of partial dissociation, motor mechanisms normally under conscious control become involuntary and automatic. Through being involuntary, they may evade the censorship to which conscious psi projections would be subjected. In a state of trance, for example, psi projections may influence the speech mechanisms or the writing mechanisms (in automatic writing) when these mechanisms have become involuntary as a result of dissociation.

We have so far been considering psi in its perceptual aspect (that is, ESP). Psychokinesis, however, is probably to be regarded as a phenomenon of the same order and it may be considered as the motor aspect of psi. Indications of this relationship between ESP and PK are to be found in the fact that both show somewhat similar peculiarities, such as terminal salience and a tendency to extinction with continued experimentation. Further evidence that they are related phenomena would be found if it were discovered that there was a correlation between success in card-guessing experiments and in the correspondence between intention and result in dice-throwing experiments. If the two phenomena are closely related, the principles of explanation adopted in the one may be expected to be adequate to the other.

We may therefore make a speculation as to PK which is parallel to the above speculation with respect to psi cognitions. Let us suppose that the primitive organism's power of movement was not

originally restricted to the organism itself. This would be a power of movement strikingly different from the kind that we are familiar with. I can now move my hand or other part of my body; I cannot move a book on the table except by establishing contact between it and some part of my body. It is possible to imagine that this power of movement might not be restricted to parts of my body and that I might be able to move the book as directly as I can move my hand. There is, of course, a real difficulty in imagining this. The difficulty is the same as that of imagining the means by which psi cognitions take place; that there is an absence of the physical continuity which we feel is necessary for the explanation of what happens. In this case, the physical continuity which our thought requires is between the object moved and the organism from which the intention originates. The fact of psychokinesis, however, has been experimentally established (3). It is necessary, therefore, to suppose that in some circumstances a movement which results from the intention of an organism may take place without this physical continuity between the organism and the object moved.

In any recognizable form, this occurs so little amongst ourselves that it is only under special conditions of experimentation that any such effect can be recognized. Under these special conditions in which psychokinetic effects are observable, it may be that we are experiencing a capacity which was more widespread at a very early stage of evolution. Such diffuse motor effects not restricted to the organism itself but extended to surrounding objects not in physical contact with it may be supposed to have been a relatively inefficient means of maintaining serviceable relations between the organism and its environment. The later development of control of movement by a motor nervous system, while restricting movements originated by the organism to the organism itself and to that part of the environment in spatio-temporal contiguity with it, may be supposed, at the same time, to have made possible an increased co-ordination between the movements of the organism and its environmental needs. In other words, the nervous control of movement and restriction of this movement to the organism itself was an evolutionary development which replaced the less efficient diffuse motility by a more efficient means of attaining control of the environment.

The function of the motor nervous system can then be regarded not only as that of transmitting impulses to the moving parts of the organism but also as a means of inhibiting the earlier diffuse motility not restricted to the organism. The occurrence of psychokinesis under experimental conditions may be considered to be the result of failure of the inhibition of this diffuse unrestricted motility, and it is to be expected that the conditions favorable to its occurrence will be those unfavorable to voluntary motor efficiency for the same reason as we expect the conditions favorable to ESP to be those unfavorable to sensory perception. The extinction of PK with continued experimentation, like the extinction of ESP in the same circumstances, is to be explained as a result of the conditions of experimentation favoring the reinforcement of the normal inhibition of diffuse motility not restricted to the organism.

The hypothesis with respect to psi phenomena which has been here sketched out is one that has the character of a scientific hypothesis in the sense that it implies certain expectations which can be experimentally disproved or verified. It does not, however, imply that the psi phenomena, either on the cognitive or the motor side, fit into the general body of scientific theory. These phenomena cannot fit into this system since current scientific theory deals exclusively with phenomena connected by a special type of continuity such as that found in sensory perception and in movement controlled by nervous impulses, glandular secretions, and so on. Psi phenomena differ from these in several important particulars: they apparently cannot be screened, they are not necessarily limited by location in a spatio-temporal frame of reference, and they seem not to require intermediate events of any known type between the cause and the effect (that is, between the cognized fact and its psi-determined response in ESP or between the intention and its physical consequence in PK). Psi cognitions are not to be regarded as being of the nature of perceptions and psi kinesis is not to be regarded as of the nature of voluntary movement, nor do they seem to belong to the system of scientific expectations within which sensory perception and voluntary bodily movement have their place.

If continuity of the type ordinarily postulated in scientific theory appears to be absent when we are dealing with psi phenomena, this

does not imply that there is here no continuity. It may be considered that continuity of some kind between an event and any consequence of that event (such as its projection on an organism's behavioral system) is a necessity of thought. The problem of stating what is the essential continuity of psi phenomena which, although transcending the ordinary limitations of space and time, yet in some way limits the primitive randomness of psi cannot yet be solved. It seems to us to be a problem for the future.

References

1. Carington, W. Experiments on the paranormal cognition of drawings. *Proc. Soc. psych. Res.*, London, 1940, **46**, 34-151; 1941, **46**, 277-344.
2. Garrett, E. J. *My Life as a Search for the Meaning of Mediumship.* New York: Creative Age Press, 1939.
3. Gibson, E. P., Gibson, L. H., and Rhine, J. B. A large series of PK tests. *J. Parapsychol.*, 1943, **7**, 228-37. (See also other papers on PK experiments in the same journal.)
4. Goldney, K. M., and Soal, S. G. Experiments in precognitive telepathy. *Proc. Soc. psych. Res.*, London, 1943, **47**, 21-150.
5. Rhine, J. B. *Extra-sensory Perception.* Boston: Bruce Humphries, 1935.
6. Rhine, J. B., Pratt, J. G., Stuart, C. E., Smith, B. M., and Greenwood, J. A. *Extra-sensory Perception after Sixty Years.* New York: Henry Holt & Co., 1940.
7. Soal, S. G. Fresh light on card guessing—some new effects. *Proc. Soc. psych. Res.*, London, 1940, **46**, 152-98.
8. Thouless, R. H. Experiments on paranormal guessing. *Brit. J. of Psych.*, 1942, **33**, 15-27.
9. ———. The present position of experimental research into telepathy and related phenomena. *Proc. Soc. psych. Res.*, London, 1942, **47**, 1-19.

83 Newmarket Road
Cambridge, England

25 Welbeck Street
London W 1, England

TWO GROUPS OF PK SUBJECTS COMPARED

By William Gatling and J. B. Rhine

ABSTRACT: A group of four young men who had been successful in the game of crap shooting were asked to participate in PK tests using dice. Four ministerial students took part in the same tests. Six dice were thrown at a time from a cup with a roughened interior. The dice were thrown an equal number of times for all of the six faces of the die. Both groups scored well above chance, giving significant results, but there was practically no difference between the average scores for the two groups. The total of 1,242 runs gives one of the highest critical ratios ever obtained in the dice-throwing tests. No conspicuous position effects were found, though the upper half of the page scored enough more hits than the lower half to give a significant difference.—Ed.

The question often arises as to whether those who are especially successful in dice games prove to be good subjects in PK tests using dice. The experiment reported in this paper grew out of a question raised by W. G. as to whether subjects who do well in the game of dice which is known as crap shooting actually possess any more PK ability than comparable individuals who have no experience in dice-throwing games. A preliminary inquiry was planned in which eight subjects, all of them young men of college age, were asked to participate in the experiment. Four of these were men who were regarded as successful at the crap game, and the other four, including W. G. himself, were ministerial students at Duke University. None of the latter had had any appreciable experience in dice-throwing games of any kind.

The experiment was carried out in the Parapsychology Laboratory at Duke University in the month of June, 1943. It extended over a period of approximately two weeks, from the fourth to the seventeenth, inclusive, and was terminated by W. G.'s departure from the University.

Procedure

The dice were thrown, six at a time from a cardboard cup with a roughened lining made of rubber, onto a standard experimental PK dice table with padded top and side walls. Three different sets of common commercial dice were used in the experiment, each con-

sisting of six dice: a set of small, white dice, 7/16 of an inch on the edge; a set of medium-sized red dice, 5/8 of an inch on the edge; and a set of medium-sized white dice 11/16 of an inch on the edge. A given set of dice was thrown exactly the same number of times for each of the six faces as the target face. This precaution was taken in order to correct for imperfections in the dice, and no further control was necessary since bias in favor of one target face would be equalized when the dice were thrown for the opposite face.

The target face for which the dice were to be thrown was entered on the record sheet at the top of each column. A column had 36 spaces, in each of which the score for a single throw of the six dice was recorded. The column was divided into three sections of twelve spaces each, but the column was scored as a whole. There were six columns to the page. The full record sheet accommodated a total of 54 runs of 24 single die readings.

The subject was allowed to choose the set of dice with which he was to work, and he was allowed to select the target face for a given column. The target face was selected and entered on the record, and the next 36 throws were recorded without omission. Toward the latter part of the experimental series, it was not possible to allow entirely free choice of target face because the number of target throws for each set of dice had to be equalized. But within the range of possibility, freedom of choice was allowed throughout the experiment.

W. G. acted as recorder at all sessions except those in which he was the subject. For the most part, the recording for W. G.'s throwing was done by another ministerial student, Mr. Paul Walenta, who also participated as a subject. Otherwise the subject who happened to be present was asked to record. The recorder called aloud the number of hits for each throw as he put it down.

The records have been independently rechecked and have been analyzed for position effects by Dr. Betty M. Humphrey of the Parapsychology Laboratory. This analysis in itself constitutes a recheck on the computation and scoring.

The selection of subjects from the ministerial students was entirely on a volunteer basis; W. G. had no preliminary knowledge of the subjects' PK ability. In the case of the four young amateur

gamblers, he had only hearsay evidence and the statements of the men themselves that they were successful in throwing dice; but the evident interest of the men in participating in the tests suggested that their belief in their capacity to influence the dice was genuine.

Results

A total of 1,242 runs were made, giving 29,808 die readings. The average score per run was 4.52, or 0.52 above the average of 4.00 expected from chance alone. There was a total of 641 hits above the mean chance expectation, which, with a standard deviation of 64.31, gives a CR of 9.97. This exceptionally large CR is an outstanding one in the field of PK investigation.

The distribution of deviation for the different target faces, each of which is represented by 207 runs, is as follows:

1's	2's	3's	4's	5's	6's
+50	+113	+136	+73	+133	+136

The four subjects with crap-shooting experience threw a total of 540 runs, with an average score of 4.52, where 4.00 is expectation. The four ministerial students threw a total of 702 runs, with an average score of 4.51. The difference is negligible. The individual subjects' data are given in Table 1.

Table 1

Results by Individual Subjects

Experienced Crapshooters				Ministerial Students			
Subject	Runs	Dev.	Av. Score	Subject	Runs	Dev.	Av. Score
A.J.	270	+215	4.80	W.G.	324	+223	4.69
H.G.	54	+ 22	4.41	C.R.	54	+ 33	4.61
V.H.	108	+ 36	4.33	P.W.	270	+105	4.39
H.B.	108	+ 9	4.08	S.A.	54	− 2	3.96
Total	540	+282	4.52	Total	702	+359	4.51

Grand Total 1242: +641±64.31 CR=9.97

The results of the analyses for position effects upon the distribution of hits present an interesting feature of the report. The principal position effect, and one of those most to be expected on

the basis of past analyses, is the decline of scoring rate on the record page, the vertical decline. The upper half of the record page scored a total of 400 hits above expectation, and the lower half, 241 hits above. This gives a significant difference of 159, with a CR of the difference of 2.47.

There is likewise a decline in the three segments of the column; that is, when the page is divided into thirds, the upper third shows a deviation of +248; the second, +246; and the third, +147.

The quarter-page distribution of hits is not the typical one, for the third quarter is slightly above the first in deviation. But the difference between the first and fourth quarters is considerable; that is, there is a noteworthy diagonal decline. A CR of the difference between the score averages of these two quarters is 2.22, which has a probability of .013. It thus falls short of the .01 criterion of significance by a small margin. These figures may be seen in the following square.



There is no chronological decline, the average scores by thirds being 4.40, 4.97, and 4.43.

Discussion

The interpretation of these results is comparatively simple. With a CR of nearly 10, there is no question about the exclusion of chance. The procedure of equalizing the number of throws for each face of the die provides a fully satisfactory provision against the possibility that faulty dice influenced the results. The manner of throwing the dice, six at a time, from a cup with roughened interior onto a padded table top would appear to insure against the operation of skills or tricks in dice throwing.

We recognize, of course, that errors in reading or recording

were a possible factor, as can always be supposed wherever human observation is involved. We have some objective assurance, however, that something happened in these experiments that could not reasonably be attributed to this general type of hypothesis, for the vertical decline already referred to produced a significant difference between the top and bottom halves of the page. There is no reason to suppose that either the reading or the recording of the dice would be less accurate at the top of the page to the extent of giving 159 more hits than the bottom. Rather, if any difference due to such errors were to be supposed, it would be most likely to occur (as a result of fatigue or boredom) in the lower right-hand quarter of the page as the session was nearing its end. But that quarter gave the lowest instead of the highest scoring.

We may with reasonable safety, then, assume that these results constitute evidence of the PK effect. If so, the first observation to be made is that they represent a relatively high average score, considering the size of the series. If any valuable clue could be obtained from the experiment as to what raised the scoring level, it would be a finding of major importance. All we can offer, however, is of suggestive nature, something perhaps for trial in further experiments.

It seems reasonable in retrospect to attribute the relatively high order of success to the fact that the experiment was a sort of playful contest. In a not-too-serious way, the crapshooters knew that they were up there in the Laboratory to show what they could do. Their "professional" pride was involved, but in a light way; it was something to joke about. The ministerial students probably felt no less the zeal to demonstrate their ability. Perhaps we might suppose they wanted to show that if they were not gamblers, it was not because they were deficient in the necessary capacities to succeed. This competitive spirit may have been the added stimulus needed to bring out better PK scoring than usual.

Moreover, the ministerial students had an appreciation of the bearing of the experimental findings in PK research. They knew that the study was directed toward a deeper understanding of human personality, and they recognized that a successful demonstration of PK would raise the level of importance of the role of the mind in

the life of the individual, disfavoring a materialistic conception of man. They were aware of the close interrelation between the objectives of parapsychology and the interests of religion. This motivation may have been of importance. At any rate, it is worth recording.

There seems to be more interest and liveliness generated in an experimental situation by an ESP or PK experimenter when he is himself first approaching the research problem. There is a vital curiosity that seems to be communicable to his subjects, probably evoking from them a more effective participation in the "game." W. G. in this experiment was doing his first work with PK, and he experienced a high degree of interest that may well have been a factor in the success obtained.

The number of subjects from each class compared in the experiment was too small to be representative. Moreover, there is no way to judge how typical of the group the four subjects were. But for whatever the comparison is worth, at best a suggestive matter, there is no reason to suppose a successful crapshooter is better in PK ability than a comparable individual from a divinity school group. As we have said, this comparison has served a sufficient and an important purpose in stimulating friendly rivalry in the tests.

131 Farnham Avenue
New Haven, Connecticut

Parapsychology Laboratory
Duke University
Durham, North Carolina

MINOR ARTICLE

HYPNOTIC SUGGESTION IN PK TESTS

By J. B. RHINE

ABSTRACT: Five young men were tested individually for PK ability by a procedure in which the subject tried to influence the fall of 96 dice at a time by means of direct mental action. After a short test session, each subject was hypnotized and told that he would try very hard to make the specified target face turn up when he released the dice and that he would be able to influence them by his concentrated effort. Although the average score was well above "chance" before the hypnosis, it dropped below in the posthypnotic test period. This drop was contrary to the effect intended by the hypnotic suggestion given. With two subjects, however, it was found that an incidental break in the hypnotic spell brought about a reversion to high scoring. Once this reversal was noted, the two subjects who thereafter scored low following the hypnosis were rehypnotized and told that they would take part in the further tests in a spirit of fun and relaxation. They too then returned to a high-scoring level, higher than their original scoring. These variations in scoring produced several significant differences. It appears clear that an effect was produced by the hypnosis, though it was not the one first anticipated on the basis of the suggestion given.—Ed.

A NUMBER of efforts have been made to increase our control over PK. In a general way, nearly every study of the conditions under which PK functions may be regarded as an effort at increasing control. Even the studies designed to interfere with the subject's success in PK tests represent contributions toward the ultimate objective of understanding and control.

And some progress has been made. On the one hand, variation of the *physical* conditions involved in the tests has so far been found to be without effect, unless indirectly through its influence upon the attitude or motivation of the subject. On the other hand, the more distinctly *mental* states of the subject have been discovered to have a very pronounced effect upon success. It was found (3) that distraction reduced PK performance to the "chance" level. The drugs, caffeine and alcohol (in strong doses), have much the same general effect upon PK success that they have upon ESP and upon higher mental functions in general—an inhibiting effect in the case of the alcohol (1), and a stimulating one in the case of caffeine (4). The

numerous studies of position effects in PK have brought out performance curves which are in some cases quite in line with similar findings in memory and learning.

The chief handicap with PK is in not knowing when it works. It is the same trouble, in fact, as with ESP: the subject is operating blindly and there is no reliable introspective observation. Therefore any effective control over PK performance would have either to bring it into the range of introspective control so that the subject could know how and when he was exercising PK, or to find some means of influencing the subject's unconscious processes. It is this latter possibility with which the present report deals in introducing the use of hypnosis into the PK experiments.

Some years previous to the experiment reported here we tried to influence performance in ESP tests by the use of hypnosis but had little success in doing so. Afterward it seemed to us likely that we had failed in the earlier effort because we did not know what suggestion to use on the subjects to assist them in improving their ESP scores.

More recently Grela (2) tried hypnosis in ESP experiments and got his highest scoring—and it was significant—when the subjects were hypnotized and given positive suggestion (suggestion intended to establish a favorable attitude and self-confidence). His lowest scoring average was obtained when subjects were given negative suggestion (suggestion intended to awaken the subject's doubt of his ability to make good scores). The latter averaged lower than the control sessions but was still above mean chance expectation. We found, however, upon analysis of Grela's data at this Laboratory, that individual differences of subjects persisted through the different states more significantly than they varied through the influence of hypnotic suggestion. Nevertheless, even though Grela did not find much hypnotic control over ESP, he did find some, and his results are therefore encouraging. And even though they concern ESP tests rather than PK, the results are worth keeping in mind in connection with the present report, which, incidentally, takes us back many years prior to the date of Grela's work.

The Experimental Procedure

It was in the spring of 1936 that a dice-throwing experiment in PK which involved hypnosis was carried out at Duke University. Five men, all of them students, served as subjects, four giving one session each to the experiment, and one contributing two sessions, making six in all. I will hereafter use these six *sessions* as the convenient units in terms of which to describe the experiment. The average session required around two hours.

This research belongs in the same group as the alcohol and caffeine work already reported. In this experiment, as in those, 96 dice were cast at a time by means of a trip arrangement which allowed them to fall by gravity from a box down an inclined plane to a padded table top having an area three by six feet. The table top was black and the dice were white, marked with black spots. The six-face was the target.

As soon as the dice box was tripped and the dice thrown, the sixes were counted by being picked up one at a time, special care being taken that other dice were not touched. In the first three sessions the subject was allowed to assist in this picking-up and counting of the dice, always, of course, under the observation of the experimenter. This participation was believed to make the test more interesting and to keep it from becoming too slow. During the last three sessions—which, incidentally, gave the more significant scoring—the subjects were not allowed to help in picking up the dice or to touch them in any way until the counting was done. At the last two sessions an assistant experimenter was present to help speed up this part of the procedure.

The same 96 dice were used as in the earlier experiments referred to. They were common commercial dice, 7/16 of an inch on the edge. The fact that the experiment was designed to involve a comparison of two different conditions in which the same dice would be used in exactly the same way made it unnecessary to attempt to obtain 96 "perfect" dice or to run any special control of an empirical nature on the ones we did use. One part of the experiment furnishes the control for the other part.

The manner in which the dice were thrown left the entire activa-

tion of them to the principle of gravitation. No kind of manual skill could enter in.

The session began with a prehypnotic control series. This consisted of 20 throws of 96 dice each. In terms of the standard run of 24 die-throws, each of these throws represented four runs. Accordingly, the tables to be presented will show 80 runs for each subject in this preliminary part of the session.

When the 80 runs were finished, the subject was seated in an easy chair. He was told to relax and to give concentrated attention to my voice. Repeated suggestions of relaxation, withdrawal from surroundings, inattention to everything except the words he was hearing, and the like soon brought about a degree of hypersuggestibility that is more or less typical of the hypnotic state. Since most of the criteria of hypnosis are artifacts incidental to the suggestion given (specific or implied) I shall make no attempt to appraise the so-called "depth" of the hypnosis. As soon as the subject was no longer able to disengage interlocked fingers when challenged to do so, to lift his feet from the floor, to open his eyes, or to speak when told that he could not, he was regarded as ready for the suggestions relative to the test.

He was then told that he would be very eager to have the dice fall with the chosen target face turned up. He was strongly assured that he would be able to keep his attention on the task of influencing the dice and that he would have great confidence in his ability to make them behave as he willed. He was then told that he would not make any effort to remember these instructions but that they would be effectively carried out. No definite suggestion of amnesia was undertaken, though in at least one case it occurred anyhow. The subject was then brought to a normally alert state and the PK experiments were resumed. The plan had been to duplicate the 20 throws or 80 runs that had been carried out before the hypnosis, but in view of developments in individual cases, this was not done exactly in every instance.

I cannot stress too strongly the importance of the fact that *hypnosis is a very individual matter.* The same words spoken to two different persons are likely to mean something different, and even if they do not, to evoke a different response. For that matter,

the experimenter himself may vary from day to day or treat different persons differently while saying the same words. We know that an experimenter changes with repetition of his role in successive experiments. Therefore, in considering the overt procedure, one must keep in mind that there is a great deal actually involved in such an experiment which cannot even be described, much less controlled.

Mr. A. J. Linzmayer, at that time secretary to the Parapsychology Laboratory, assisted in counting and recording in two of the experimental sessions. The records have been completely rechecked and the mathematical work independently computed by Dr. Betty M. Humphrey also of the Parapsychology Laboratory.

Results in Terms of Individual Sessions

The experiments did not turn out as had been expected. In view of that fact and the fact that the entire research was still in an exploratory stage, some modifications were introduced into the plan of the experiment as it was being carried out. The explanation will be clearest if the results and variations are introduced by sessions, with a summary and general treatment at the end.

Subject W.R.A.

In the preliminary 80 runs W.R.A. obtained a deviation of +15, which is an average of 4.19. After the hypnosis his 80 runs gave a deviation of −30 and a very low average of 3.62. This result was just the opposite of what had been expected, and the drop from +15 to −30 is almost significant. It indicated that the suggestion had operated upon the subject's PK capacity in some reverse manner. I asked him to wait while I worked with the next subject in order to continue the tests afterward, for I was most interested. It was at least an hour later when I got back to him, and without any further hypnosis he made 10 more throws, or 40 runs. From these he obtained a deviation of +14, with an average score of 4.35, as compared with his preliminary average of 4.19. He was presumably no longer under the effect of the hypnotic suggestion during these last 40 runs. In fact, he had been a witness to the performance of subject C.D.C. and acted as if he wanted to see what *he* could do in comparison.

Subject C.D.C.

The preliminary 80 runs with subject C.D.C. gave a deviation of +4 and an average score of 4.05. After hypnosis the first 80 runs gave +23, or 4.29 as the average. This was in the direction expected. Since I was going to do an additional short series with W.R.A., who was waiting, and since subject C.D.C. seemed interested in going on, I had him continue for another 40 runs without a break. These gave a deviation of +18. The average score of this series went up higher still to 4.34. (Then we went back to subject W.R.A. and he did his 40 runs with the average score of 4.35.)

Subject C.E.W.

C.E.W. started off low with a deviation of −6 on his first 80 runs, an average of 3.92. After hypnosis, he dropped still lower and, in the first 72 runs, averaged 3.85, a deviation of −11. At this point in the series, this subject, who had been averaging below mean chance expectation from the beginning of his session, asked nervously if he could smoke a cigarette. His manner of asking and his demeanor suggested tension. After he had smoked the cigarette, a marked change came over him both in terms of appearance and in the results of his tests. The difference was so marked that he was kept at the task until he had completed 88 runs, eight which the plan of the experiment had called for and 80 more to test this new condition which the change accompanying the cigarette seemed to have induced in him. The 88 runs after the cigarette showed a deviation of +32, an average of 4.36.

The performance of C.E.W., following upon that of W.R.A., made it look very much as if the hypnotic suggestion that was being given was certainly not of the right kind. At least it was giving us the wrong type of effect. The two subjects together had a deviation of −41 on 152 runs right after hypnosis, but after a time-break in the case of W.R.A. and the break of smoking a cigarette in the case of C.E.W., the two of them together scored a deviation of +46 in 128 runs. This did not look accidental. There was, in fact, a significant difference in scoring rate between the two stages. The CR of the difference is 2.90. At this point, therefore, a change in plan

was introduced: if subjects continued to give a lower scoring rate with hypnotic suggestion, they were to be rehypnotized and given a different type of suggestion.

Subject D.V.H.

In his first 80 runs D.V.H. scored 4.26, a total deviation of +21. After hypnosis his 80 runs gave exactly the opposite, −21, and an average of 3.74. Thereupon the subject was rehypnotized and told that now he would be relaxed, feel free and easy about the tests, and would throw the dice more in the spirit of a game which he would enjoy. Following this second hypnotic session, he began at once to score on the positive side of the mean. He was kept going through 140 runs, ending with a deviation of +41, an average score of 4.29. This result definitely favored the view that I had been giving the wrong kind of suggestion.

Subject C.D.C., Second Session

In his second session C.D.C. started off with his first 80 runs on the highest level he had reached; namely, an average score of 4.44, or a deviation of +35. But after hypnosis this time he dropped to a negative deviation. In the first 15 throws, or 60 runs, he gave a deviation of −18, with a score of 3.70. Whatever the reason, he complained of a headache. It may have been the result of strain and conflict induced by the effort to score and his disappointment with the low scoring. At that point (after 15 throws) he asked me if I could do something to relieve his headache. I rehypnotized him and told him that he would feel well again immediately and would enjoy continuing the tests. I went on to give him essentially the same suggestion as I had given D.V.H. in his second hypnotic treatment. Following the second hypnotization, C.D.C. at once obtained scores above expectation. He made only seven throws, or 28 runs, before the experiment was interrupted, but in that time he accumulated a positive deviation of 8 and an average score of 4.29.

Subject L.E.S.

In his first 80 runs L.E.S. obtained a deviation of +22 and an average score of 4.28. After hypnosis he did better still. Eighty

runs gave +36, an average score of 4.45. L.E.S. had not wanted to be hypnotized in the first place, and I thought it best not to urge him to submit to being rehypnotized.

General Results

The results of all six sessions are summarized in Table 1 in the order in which I have discussed them. The prehypnotic results make up Section A. The results taken after the first hypnotic treatment and up to the point where the series was broken comprise Section B. All of the tests that came after the break are grouped

Table 1

The PK Results by Sessions, Showing Prehypnotic and Posthypnotic Results

Subject	Section A (Prehypnotic)			Section B (First Posthypnotic)			Section C (Miscellaneous)		
	Runs	Dev.	Av. Score	Runs	Dev.	Av. Score	Runs	Dev.	Av. Score
							(After a break)		
W.R.A.	80	+15	4.19	80	−30	3.62	40	+14	4.35
C.D.C.	80	+ 4	4.05	120	+41	4.34			
C.E.W.	80	− 6	3.92	72	−11	3.85	88	+32	4.36
							(Second posthypnotic)		
D.V.H.	80	+21	4.26	80	−21	3.74	140	+41	4.29
C.D.C.	80	+35	4.44	60	−18	3.70	28	+ 8	4.29
L.E.S.	80	+22	4.28	80	+36	4.45			
Total	480	+91	4.19	492	− 3	3.99	296	+95	4.32

Total runs = 1268; Av. score = 4.14.

together as Section C. This section includes the two subjects, D.V.H. and C.D.C., who were rehypnotized, as well as C.E.W.'s results after the cigarette and W.R.A.'s 40 runs after the interruption.

The data of the period immediately following hypnosis are appropriately grouped together since they were done before any change was introduced. They give a total of 492 runs, a negative deviation of −3, and an average score of 3.99. The Section A total is 480 runs; deviation, +91; average score, 4.19. The four subjects who contributed to Section C give the highest average score, 4.32 for 296 runs.

Some of the more general relationships brought out by the analysis of these data are presented in Table 2. The data of Section B, those obtained immediately after the first hypnotic treatment, differ significantly from those of Sections A and C combined, giving a CR of the difference of 2.34. Although this CR would

Table 2

DIFFERENCE IN SUCCESS BETWEEN FIRST POSTHYPNOTIC AND OTHER TESTS

Subject	Sections A and C		Section B		CR_d	x^2
	Runs	Av. Score	Runs	Av. Score		
W.R.A.....	120	4.24	80	3.62	2.38	5.66
C.D.C.....	80	4.05	120	4.34	1.12	1.25
C.E.W.....	168	4.15	72	3.85	1.15	1.32
D.V.H.....	220	4.28	80	3.74	2.25	5.06
C.D.C.....	108	4.40	60	3.70	2.41	5.81
L.E.S.....	80	4.28	80	4.45	.61	.37
Total.....	776	4.24	492	3.99	2.34	$\Sigma x^2 = 19.47$ P = .004

ordinarily be regarded as significant ($P = .01$) there is some correction needed here for the fact that other differences would have been of research interest as well as the one selected.

The striking individual differences in behavior of the subjects suggest, however, that pooling obscures the observed changes. Accordingly a more satisfactory measure can be had by taking the differences between Section B and the pooled data of A and C together by individual sessions. The CR's of the difference are given in Table 2, and these are combined by the chi-square method. The chi-square of 19.47 for six cases has a probability of .004. Not only is this significant; it could also stand some correction if needed. Perhaps the only correction to be considered is whether Section C might not have been pooled with B instead of with A. This would lead to a multiplication of the P of .004 by two, giving .008, which is still significant.

The significant chi-square based on differences in either direction for the individual sessions cannot be explained by the hypothesis of

faulty dice.[1] It particularly focuses interest on the research objective; namely, differential performance as a consequence of hypnotic suggestion.

Discussion

Consideration of Alternative Hypotheses

The work reported in the preceding section is best explained by the PK hypothesis. The *differences* shown by sessions between Section B and the sum of Sections A and C are significant when combined by the chi-square method. This rules out the chance hypothesis regardless of the trueness of the dice. We need not attempt to determine just how free from imperfection the dice were, for it is the comparative data that are important here. Moreover, the uniform, semimechanical way in which the dice were thrown allowed no opportunity for the use of skills. The conclusions rest entirely upon the basis of the difference between conditions.

Differences of judgment might arise as to the propriety of grouping in Section C the four batches of data that are found there. There can hardly be any question about the wisdom of a separate grouping for the data of subjects D.V.H. and C.D.C. who were rehypnotized; but the cases of W.R.A. and C.E.W. could be considered as not properly belonging with this rehypnosis group. The point is debatable. As I have explained, they are included in Section C because it seems reasonable to assume that what happened in each case broke the spell of the posthypnotic suggestion operative in producing the data of Section B. W.R.A.'s long delay in waiting for C.D.C. to perform certainly puts his Section C data in a different category from those he did in Section B. What the effect of the cigarette was on C.E.W. we can only infer, and that in a tentative way. The mere fact that the subject got up enough independent motivation to ask for a cigarette, his evident nervousness,

[1] The CR of the difference on which this computation was made should ideally be recomputed by the arc sine transformation method to correct for bias in the dice as proposed by Dr. J. A. Greenwood in his article in the Journal of Parapsychology for December, 1943. But as readers who have followed the recent PK reports are aware, these corrections have invariably been slight and insignificant. The considerable labor of computing the six cases required for the data of Table 2 does not therefore appear justified in this instance. It can be done by anyone who wishes to do so on the basis of the published method referred to and the data provided by the table.

and the signs of marked relaxation and general well-being after the smoking definitely suggest to the experimenter that the effect of the posthypnotic suggestion was altered and that the subject had recovered something he did not have earlier in his Section B throwing.

Perhaps it is best simply to regard Section C as a miscellaneous grouping. Yet all four of the sessions represented there have this much in common: in them the subjects had all reacted by giving negative deviations in Section B, and they all had had something happen which could reasonably be taken as releasing them from the posthypnotic suggestion. The fact is that they all scored well above expectation after the break. Their scores thereafter differ to a significant degree, not only from their own scores in Section B, but also from the scores of Section B as a whole (that is, all six sessions). These facts make it clear that something distinctive had happened and leave no question, I think, but that the effect of the first hypnosis upon these subjects had terminated.

Someone may wish to consider apart from the rest the cases of those two subjects, D.V.H. and C.D.C., who were hypnotized a second time. This step leads, however, to the same conclusions since these two subjects together gave the following results:

	Section A	Section B	Section C
Runs	160	140	168
Deviation	+56	—39	+49

The difference between Sections A and B is significant (CR_d = 2.99); so also is that between B and C (CR_d = 2.74). Finally, Section B differs from the sum of the other two to a significant degree (CR_d = 3.25).

Before the matter of counterhypotheses is dismissed, a word or two may be said about other possibilities. In dealing with large numbers of dice, the chance of error in counting is recognized to be greater than it is with small numbers. As has been pointed out before, this danger is almost certain to be on the side of omission; that is, in counting the number of hits, experience has taught us that we may overlook a die showing the target face. Danger of the other type of error—of overcounting—is very slight. The die has to be picked up, set aside, and then, at the end, recounted. It seems

fairly certain that, in this experiment, counting errors could not have produced the results on which the conclusions are based, since it is the difference between the sections rather than the total score upon which the focus of attention is directed. So far as motivated errors are concerned, the mere fact that Section B had been anticipated as likely to give high scores is probably a sufficient answer on that question. There is no selection involved, for all of the work done under these conditions is included here.

We may proceed, then, to the discussion of the results with the understanding that none of the familiar counterhypotheses apply to the results contained in this report. The conclusion is that we are dealing with the PK process.

Position Effects

In the 96-dice series, of which this paper is the third, the data were recorded in short columns of five entries each, each entry representing the number of hits made in a throw of the 96 dice. The two earlier reports, the alcohol and caffeine papers, presented, in a total of 1,120 runs, the following distribution of hits in terms of the deviations from mean chance expectation at each of the five places in the column. These deviations include the results of the control tests made before the drug was given and those made after it was given. Both the "before" and "after" tests as a whole show the decline, and both drug experiments contributed to it.

Place in the column	1	2	3	4	5
Deviation	+100	+127	+50	+27	+16

These show a very marked decline from the first trial to the fifth, though the second was better than the first. In the hypnosis experiment, although the first trial was the highest, there was no great amount of decline at the end of the column; rather, there was a considerable upturn, giving a somewhat U-shaped curve. There is no pattern of special interest in the distribution of any section of the data, and only the pooled deviations are given for the column.

Place in the column	1	2	3	4	5
Deviation	+54	—11	+9	+40	+23

There was a difference in the type of subjects in the two earlier series which may have been responsible for the difference in distri-

bution. The alcohol and caffeine experiments were done with experienced subjects; the hypnosis experiments were not. An experienced subject is more likely to know what the structure of the run is and to keep track of the progress in terms of the record sheet. The naive and beginning subject does not, for a time at least, structure the task in which he is participating. Each throw is more of a separate event, for it is all new to him. The most significant position effects found in the PK research thus far have been found in the results of experiments in which one subject, or two, operated through long periods either in doing their own recording or in sharing in the record-taking, so that the structure of the test units was deeply ingrained upon their attention.

Number of Dice per Throw

We have again, in this series, confirmation of the fact that successful PK performance can be achieved with the throwing of 96 dice at a time. Comparisons cannot profitably be made with other series in which smaller numbers of dice were thrown at a time because the dice, target, and other features were dissimilar. But for what it is worth, an overall perspective of the published work suggests a generally favorable comparison between the scores made with 96 dice per throw and those with the smaller numbers of dice. The simple fact that significant scoring has repeatedly been produced with the larger number of dice suggests that the number of dice per throw is not a determining factor in PK tests.

The Role of Hypnosis

The results that are of primary interest have to do with hypnosis. Regarded superficially, they were disappointing; that is, hypnosis was not the successful controlling instrument in these PK experiments that it has been found to be in a number of other areas of human ability; but as I explained in the introductory section, there are many considerations that make this somewhat understandable.

The important finding, and one that stands out fairly clearly, is that there *was* an effect. The hypnosis did something, even though in four out of the six cases it was a reversal of the intended effect. The fact that hypnosis was doing something received confirmation from the rehypnosis results of subjects D.V.H. and C.D.C. Their

differences alone, it may be remembered, were quite significant. As a beginning step, then, the experiment *was* successful. We need only compare it to the results obtained from the early Duke experiments in ESP in which hypnosis was used for a time in certain experiments. There we got no important difference between the hypnotic and nonhypnotic tests. It would have been most illuminating if we had found any significant differences whatever as a consequence of the hypnosis.

Why did four of the subjects drop in their scoring rate after hypnosis and even go so far below the theoretical mean as to give a significant negative deviation? The first answer likely to come to mind is that perhaps they were negativistic subjects. One of them, however, C.D.C., had responded earlier to hypnotic suggestion in the direction desired. Only one of the four subjects who dropped in scoring after hypnosis showed any negativistic signs in the process of hypnotization. But another subject, L.E.S., who showed similar negative signs and hesitations while being hypnotized, raised his score average after hypnosis. I doubt, therefore, whether any general negativistic tendency can properly be evoked as an explanation. Hypnosis by positive suggestion would not likely have been successful with any appreciable negativism present.

Probably there *was* conflict and strain induced in these subjects. A suggestion of strong concentration probably brings into the subject a state of general tension that is inhibitory to the process of PK. In other words, the subject tries very hard in the wrong way, inducing concentration of functions that have nothing to do with PK.

Why these four subjects went so far below expectation instead of simply dropping to that level is another difficult question. It looks as if they were doing something to the dice but doing the very opposite to what they were consciously attempting. Here again we must recall that the essential act of PK is wholly unconscious; so what it was that the subjects were wanting to do on the level of unconsciousness where PK actually occurs would be difficult to discover. Quite conceivably, something in the act of the first hypnotizing may have antagonized the subjects. Every hypnotist will recognize that very often a subject will show facial or other signs

of serious conflict, indicating that there is a strong component of resistance operating in his mental organization somewhere. Usually it is unexpressed except in indirect fashion. Perhaps these PK subjects unconsciously resisted the more serious and determined way in which the hypnotist went after them in the first treatment and were much more comfortable after they were released from this, as in the second hypnotic treatment.

The most encouraging thing that stands out is that PK can be manipulated, even though it is unconscious. We know, of course, that the subject himself consciously orients the performance, deliberately wills that the target face of the die will come up, and has, in fact, everything about the operation under conscious control except awareness of the PK function. Here introspection is "stone blind." If we keep in mind this cardinal handicap in the appraisal of such results as we have, we find a great deal more encouragement and significance in the findings.

References

1. Averill, R. L., and Rhine, J. B. The effect of alcohol upon performance in PK tests. *J. Parapsychol,* 1945, 9, 32-41.
2. Grela, J. J. Effect on ESP scoring of hypnotically induced attitudes. *J. Parapsychol.,* 1945, 9, 194-202.
3. Price, M. M., and Rhine, J. B. The subject-experimenter relation in the PK test. *J. Parapsychol.,* 1944, 8, 177-86.
4. Rhine, J. B., Humphrey, B. M., and Averill, R. L. An exploratory experiment on the effect of caffeine upon performance in PK tests. *J. Parapsychol.,* 1945, 9, 80-91.

Parapsychology Laboratory
Duke University
Durham, North Carolina

BOOK REVIEW

TELEPATHY: AN OUTLINE OF ITS FACTS, THEORY, AND IMPLICATIONS. By Whately Carington. London: Methuen & Co., Ltd., 1945. Pp. 176.[1]

Our readers are already familiar with various aspects of Whately Carington's experimental work with free drawings, and with the central theoretical interpretation which he offers for all his results. The reviewer must choose between (1) a bird's-eye view of Carington's research and theoretical work as embodied in this volume, and (2) an account of those features of the present volume which go beyond earlier publications. Since the second method would involve mutilation of a well-integrated approach, the indulgent reader will have to be asked to forgive our following the former method.

"Part I: Facts" includes chapters on the history of the study of spontaneous cases of telepathy and discusses simply and clearly the early experiments, such as those of the Groningen experimenters and of Estabrooks, and the recent work of the Duke University group and of Tyrrell and Hettinger. Carington's own experiments and the method of assessing their results naturally receive much attention. This part of the book concludes with reference to the Soal-Goldney work on precognition and with a discussion of the difficulties of fitting all these data into the usual physical frames of reference. This historical and critical section is highly competent and perceptive and presents nothing which is not of permanent interest to the serious student. It makes abundantly clear the necessity for rigid experimental controls and competent statistical evaluations. It shows that when targets are suitably randomized, free drawings may be used by the percipient and assessed by a catalogue method in which the value of a hit is determined by the rareness with which each such response is given by a large percipient group.

"Part II: Theory" briefly shows the untenability of the ordinary

[1] Since the review was written, this book has appeared in an American edition *Thought Transference*, published by Creative Age Press. Pp. x + 287. $2.50.

physical theories and then proceeds at length with Carington's "association theory." The hypothesis is extremely simple: If the agent has in mind two ideas which become associated one with the other and if the percipient has in mind one of these two ideas, the fact that they are associated in the agent's mind will make more likely the appearance of the same association in the percipient's mind. Stated more generally, "If two ideas, A and B, are presented together or in close succession to any mind, M, and if subsequently one of them be re-presented to that mind *or presented to any other mind*, M', then the other idea is more likely to accompany or closely follow it in the mind to which it is so presented or re-presented than if the two ideas had not been presented together to the mind, M." This is of course a form of the "one-big-mind theory" and has in particular a good deal in common with Myers' conception that despite our surface individuality we are all conjoined in one reality at a deeper level.

It is easy, in the light of the theory, to provide a common core of experiences for agent and percipient. They may both look at the same kind of object—say two prints of the same photograph—and whatever the agent associates with the photograph should tend to come also into the percipient's mind. Such a common core of experience, or "K-object," may indeed be the "idea of the experiment" itself.

One great merit of the theory is that one may draw from it a series of corollaries, each of which may be experimentally tested. We may go ahead with all the classical laws of association and find whether inter-individual association obeys the same laws as strictly personal association. We learn, for example, that in many of Carington's experiments the drawings which were to be used had been both listed and drawn beforehand so that associative links were strong. In other series they were listed and drawn but not used. The quantitative results fit well with the theory. "The average scores were: Listed, drawn, and used, 21.2; listed and drawn, but not used, 4.3; listed only and neither drawn nor used, 2.2; controls, 0.2."

There follows an application of the theory to cases of spontaneous telepathy and to the whole realm of apparitions, showing ingeniously the role of K-ideas in telepathic phenomena of all sorts.

The theory undertakes to explain telepathy but not clairvoyance, and those who believe that a theory must be contrived which will simultaneously explain both will be dissatisfied. There is only a meager and (at least for this reviewer) vague discussion of precognitive telepathy, the time dimension being left, as it usually is, without essential clarification. For such sins of omission it would be easy to belabor the author. The physical phenomena of psychical research are left out of account altogether. The reviewer feels on the other hand that the attempt to solve all problems at once is a besetting sin, not only with philosophers but with psychical researchers and with most psychologists, and that we ought to be very profoundly grateful for so lucid and so ingenious a theory. The theory has been sifted and redefined over and over again and tested very laboriously with the ingenious catalogue method in many experiments both of Carington's own and of other investigators; and it stands up extraordinarily well. This alone would make it a major landmark in the development of our subject.

It is true, as Carington is well aware, that as far as telepathy is concerned the theory explains too much. Throughout the history of association psychology, the question has always been asked: What keeps associative links from working infallibly? What causes interference—forgetting, dissociation, failure of unity? In the same way the reader will ask here why the effects of association are so slight; why so many oft-repeated associations fail to appear in the percipient's mind; why with the universe of human experience, past and present, to draw upon, children must arduously struggle to learn such things as their ABC's. We know, of course, that Carington's theory must move here in the direction of a theory of interference; associations must fail because crowded out by other associations. Far more work is imperative if this explanation is to make good. But a theory which is so clear and workable that it shows, as this one does, where the weak spots lie and how they can be corrected is a very exceptional theory. It may be felt that there is a certain paranoid effort here to make everything fit a simple hypothesis; but it would be fair to answer that the author sprinkles his pages with constant reminders as to the tentative character of suggestions offered.

"Part III: Implications" is an attempt to deal with various questions of psychology, of psychical research, and of philosophy by an expansion of the association theory. Elementary sense impressions and their derivatives and compounds are called psychons; a cohering group of psychons is a "psychon system." Multiple personalities are regarded as psychon systems derived from one larger system, and group psychology is conceived in terms of the sharing of K-ideas resulting in a group psychon system. The survival problem is conceived as a problem in the capacity of a well-integrated psychon system to maintain itself beyond death, and indeed quite possibly to maintain itself indefinitely. These are charming speculations which in many aspects are very remote from any possibility of verification. In the reviewer's judgment they are legitimate and intriguing but in a very different universe of discourse from Carington's very substantial experiments and his magnificently clear and workable association theory of telepathy.

It may of course be argued that we have too much theory, too few facts. It is possible, however, to come to a very different conclusion; namely, that the insufferable chaos in psychical research is largely due to the absence of a genuine theoretical system such as the atomic theory in chemistry, or the evolutionary theory in biology. It is too soon to be sure, but the reviewer thinks that the Carington theory may well prove, with modifications, to be the twentieth-century atomic theory of psychical research. Ultimately it will be superseded, but in the meantime it may well prove to be a major boon to all serious experimental work.

GARDNER MURPHY.

GLOSSARY

In order to avoid constant redefining of commonly recurring terms in papers appearing in this JOURNAL, the following definitions are submitted for convenient reference.

*For a simple description of those terms marked by an asterisk, as they apply to the ESP test data, see Chapter VIII and the Appendix of *A Handbook for Testing Extra-Sensory Perception* by C. E. Stuart and J. G. Pratt. Further explanation may be found in any elementary statistical text.

AGENT: In tests for telepathy, the person whose mental states are to be apprehended by the percipient. In GESP tests, the person who looks at the stimulus object.

AVERAGE SCORE: Average number of hits per run.

CALL v.: To attempt to identify by ESP a stimulus object or a mental state of an agent.

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

*CHANCE: The complex of undefined causal factors irrelevant to the purpose at hand.
CHANCE EXPECTATION = MEAN CHANCE EXPECTATION: The most likely score if only chance obtains.
CHANCE AVERAGE: Mean chance expectation expressed as an expected score, generally in terms of average per run.

CHI-SQUARE: A sum of quantities each of which is a deviation squared divided by an expected value. Also a sum of the squares of CR's.

CLAIRVOYANCE: Extrasensory perception of objective events as distinguished from telepathic perception of the mental state of another person.

*CR (CRITICAL RATIO): A measure to determine whether or not the observed deviation is significantly greater than the expected random fluctuation about the average. The CR is obtained by dividing the observed deviation by the standard deviation. (The probability of a given CR may be obtained by consulting tables of the probability integral, such as Pearson's.)
CR OF THE DIFFERENCE: The observed difference between the average scores of two samples of data divided by the standard deviation of the difference.

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

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

DIE THROW: The elementary unit of data in a dice-throwing test represented by the throwing and reading of a single die, no matter how many dice are thrown at a time.

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

ESP (EXTRASENSORY PERCEPTION): Response to an external event not presented to any known sense.

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

ESP SYMBOLS: See plate opposite page 1, this JOURNAL, Vol. I, March, 1937.

EXPECTATION; see CHANCE.

EXTRACHANCE: Not due to chance alone.

FREE MATERIAL: The stimulus objects of experiments in which an unlimited or unspecified range of stimulus objects is employed (as contrasted with methods such as card-calling in which the subject knows that the stimulus object is one of a known range).

FREE RESPONSE METHOD: An ESP test method in which a relatively unlimited range of stimulus objects is used and in which the subject is instructed to express freely as his response any range of thoughts and ideas.

GESP (GENERAL EXTRASENSORY PERCEPTION): A technique designed to test the occurrence of extrasensory perception, permitting either telepathy or clairvoyance or both to operate.

HIGH-DICE TEST: A PK technique in which the aim of the subject is to try to influence a pair of dice to fall with the two upper faces totaling eight or more.

LOW-DICE TEST: A PK technique in which the aim of the subject is to try to influence a pair of dice to fall with the two upper faces totaling six or less.

MEAN CHANCE EXPECTATION; see CHANCE.

*P (PROBABILITY): A mathematical estimate of the expected relative frequency of a given event if chance alone were operative.

PARAPSYCHOLOGY: A division of psychology dealing with those psychical effects which appear not to fall within the scope of what is at present recognized law.

PERCIPIENT: The person who makes the calls in an ESP test.

PK (PSYCHOKINESIS): The direct influence exerted on a physical system by a subject without any known intermediate physical energy or instrumentation.

PRECOGNITION: Cognition of a future event which could not be known through rational inference.

PREFERENTIAL MATCHING: A method of scoring free responses. A judge ranks the stimulus objects (usually in sets of four) with respect to their similarity to, or association with, each response; and/or he ranks the responses with respect to their similarity to, or association with, each stimulus object.

PSI: A general term to identify personal factors or processes in nature which transcend accepted laws. It approximates the popular use of the word "psychic" and the technical one, "parapsychical."

PSI PHENOMENA: Occurrences which result from the operation of psi. They include the phenomena of both ESP (including precognition) and PK.

QD (QUARTER DISTRIBUTION): The distribution of hits in the record page (or in a logical subdivision thereof, such as the set or the half-set) as found in the four equal quarters formed by dividing the selected unit horizontally and vertically.

RUN: A group of trials, usually the successive calling of a deck of 25 ESP cards or symbols. In PK tests, 24 single die throws regardless of the number of dice thrown at the same time.

SCORE: The number of hits made in one run.

TOTAL SCORE: Total of scores made in a given number of runs.

AVERAGE SCORE: Total score divided by number of runs.

*SD (STANDARD DEVIATION): Usually the theoretical root mean square of the deviations. It is obtained from the formula $\sqrt{npq}$, in which n is the number of single trials, p the probability of success per trial, and q the probability of failure. (For ESP cards, SD $= 2\sqrt{\text{no. of runs.}}$) Sometimes the SD used is that estimated from the observed variability in the scores.

SD OF THE DIFFERENCE: For both ESP cards and PK tests using dice, the SD of the difference is equal to $\sigma_s \sqrt{1/R_1 + 1/R_2}$ where σ_s is the SD of a single run and R_1 and R_2 are the number of the runs in the respective samples compared. This gives the SD of the difference for run score averages.

SERIES: Several runs or experimental sessions that are grouped in accordance with a stated principle.

SESSION: A unit of an ESP or PK experiment comprising all the trials of one test occasion.

SET: A subdivision of the record page serving as a scoring unit for a consecutive group of trials, usually for the same target.

SEVENS TEST: A PK technique in which the aim of the subject is to try to influence a pair of dice to fall with the two upper faces totaling seven.

*SIGNIFICANCE: A numerical result is significant when it equals or surpasses some criterion of degree of chance improbability. Common criteria are: a probability value of .01 or less, or a deviation in the expected direction such that the critical ratio is 2.33 or greater.

SINGLES TEST: A PK technique in which the aim of the subject is to try to influence dice to fall with a specified face up.

STIMULUS OBJECT: The ESP card or drawing or other object, some identifying characteristic of which is to be apprehended by the percipient.

SUBJECT: The person who is experimented upon. In ESP tests, most commonly the percipient (though also the agent in GESP and telepathy). In PK tests, any individual whose task it is to influence the objects thrown.

TARGET: In ESP tests, the stimulus object; or in telepathy, the mental state of the agent. In PK tests, the faces of the die (or combination of faces) which the subject attempts to bring up in the act of throwing.

TARGET CARD: The card which the percipient is attempting to perceive (i.e., to identify or otherwise indicate a knowledge of).
TARGET DECK: The deck of cards the order of which the subject is attempting to identify.
TARGET FACE: The face on the die which the subject tries to turn up as a consequence of direct mental action.

TELEPATHY: Extrasensory perception of the mental activities of another person. It does not include the clairvoyant perception of objective events.

TRIAL: In ESP tests, a single attempt to identify a stimulus object. In PK tests, a single throw of the dice or other objects thrown.