

BF
1001
.J86

The Journal of Parapsychology

A SCIENTIFIC QUARTERLY DEALING WITH EXTRA-SENSORY
PERCEPTION AND RELATED TOPICS

Volume 6

MARCH, 1942

Number 1

Contents

| | PAGE |
|---|------|
| Editorial: The Change in Editorship and the New Program | 1 |
| Five Years of the Journal of Parapsychology | 5 |
| DOROTHY H. POPE and J. G. PRATT | |
| An ESP Test With Drawings | 20 |
| C. E. STUART | |
| A Check on Saliency Relations in ESP Data | 44 |
| J. G. PRATT, BETTY M. HUMPHREY and J. B. RHINE | |
| A Transoceanic ESP Experiment | 52 |
| J. B. RHINE and BETTY M. HUMPHREY | |
| Glossary | 75 |

DUKE UNIVERSITY PRESS
DURHAM, N. C.

The Journal of Parapsychology

Editors

J. B. RHINE

C. E. STUART

J. G. PRATT

J. A. GREENWOOD

Statistical Editor

DOROTHY H. POPE

Managing Editor

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

Contributions submitted for publication and all editorial communications should be addressed to the Managing Editor, Dorothy H. Pope, College Station, Durham, North Carolina. Correspondence with the editors is advised before submitting articles other than reports of experimentation. Since it does not forward manuscripts by registered mail, Duke University Press cannot guarantee that they will not be lost in transit, and contributors are urged to keep copies of their papers. All contributions should be typewritten, double spaced. References should be given in the form adopted by THE JOURNAL.

Reprints may be ordered when the proof is returned. THE JOURNAL will bear one half the cost of reprints up to two hundred copies. Correspondence concerning subscriptions, change of address, back numbers, and other business communications should be addressed to the Parapsychology Laboratory, College Station, Durham, N. C.

The subscription price is \$4.00 a year; single current numbers \$1.00. The rate for back volumes is \$5.00; for single numbers \$1.25. Missing numbers will be supplied free when lost in the mails if written notice is given within one month of the date of issue. All remittances should be made payable to the Parapsychology Laboratory.

Entered as second class matter at the post office at Durham, N. C.

DUKE UNIVERSITY PRESS

The Journal of Parapsychology

Volume 6

MARCH, 1942

Number 1

AN EDITORIAL:

THE CHANGE IN EDITORSHIP AND THE NEW PROGRAM

FOR THE last three of the five years that have passed since the launching of this journal, Drs. Gardner Murphy and Bernard F. Riess have borne the responsibilities of editing it. They generously did so in order that those upon whom the task would otherwise have fallen might devote themselves more fully to research in the field concerned. But because of the complications which the war and its parallel developments have introduced for them, Drs. Murphy and Riess have now asked to be relieved; and in appreciation of their service to the field, we could hardly do less than take our turn and assume the editorial burden.

We do not, however, accept this role of editorship at all lightly, in this of all times. And it is not without considerable concern for the future of the JOURNAL that we now take over—a concern that no doubt reflects something of the general anxiety which all Americans must feel today for the future of the very freedom to choose one's field of endeavor.

We enter upon the undertaking, therefore, clearly resolved that we will carry on this phase of our work, as indeed the research itself, only because we believe it can be done without preventing anyone concerned from lending all possible aid to insure victory in the present military struggle—that victory which must come before the right to do such research, or any other cherished human right, is safe. This naturally leaves many uncertainties for an editorial program, but we are determined to keep both JOURNAL and research alive through-

out the troubled years ahead—provided it can be done under the conditions just stated.

If we succeed, it will be in great part because the cooperation of the friends of parapsychology will have made it possible. The work will of necessity be a joint endeavor.

* * * *

Many changes in the *JOURNAL* are either introduced in this number or are contemplated for those to follow, but in no instance may these be properly construed as a reflection upon our predecessors. These changes will not in the main need to be justified here. The reasons for their introduction will be sufficiently evident, as they appear in this and in future issues, and the reader's judgment of their value will be genuinely welcomed.

But a few introductory words are in order concerning certain of the innovations made—for one example, the discontinuance of the Board of Review. The membership of this board had, before the change of editorship occurred, been reduced by resignation to half its usual size. The former editors found the plan unsatisfactory, especially with respect to the delay in publication caused by the reviewing of articles in the manuscript stage. The original intent in the appointment of the Board was to have it comment on reports as they appeared in print; however, for that purpose we obviously need not limit criticism to a specified number of persons—all comments by any reader should, we believe, be given careful attention and the mark of special authority reserved for none. If such comment by the reader raises questions not already answered fully in the article concerned, the matter will be referred to the author for reply, and both the question and the reply will appear in the next issue in the column headed "Notes and Correspondence." The *JOURNAL* has always welcomed criticism, and it is to be hoped that all critical readers will accept our cordial invitation to participate in the discussion of the research reports as they are published. Particularly is it now in order to emphasize the need for constructive suggestions, for additional implications that may have been overlooked by the author, and for interpretations of the findings. We trust that this plan will secure for the *JOURNAL* the real editorial benefits which had been envisaged in establishing the Board of Review system.

General Library
Psychology
Subject
10-24-46
5-6149
v. 8, no. 1-4

Such changes as the return to regular printing and the use of larger type are features which we feel confident will, in the long run, more than repay the increase in cost. The advance in subscription rate is due only in part to these innovations. Probably it can be reduced after the present national emergency is past. We frankly admit that there is a venturesome aspect to some of these features which we hope will be justified by events.

Another, but not a radical, departure will be the introduction of a wider range of subject matter and also of mode of presentation. The usual type of research report will continue to be the backbone of the JOURNAL. But we think there is now a need for other types of articles and reviews, and that those who want primarily to read the one will in many instances read the other too. So they both belong in the same organ of publication. The "general" articles will presume less acquaintance with the special terminology of this field and with scientific and statistical method; indeed, the average intelligent layman or the college undergraduate should be able to understand them without difficulty.

Articles of more technical nature will still be prefaced by abstracts, and the plan is to make these synopses somewhat fuller and less technical than heretofore, with the object of giving the essence of the research reports in general language. The glossary in the back will help the reader who approaches the technical article without previous experience with the special literature of the field.

* * * *

Five years in the life of a scientific periodical is not long, but it is sufficient to afford a view of trends as well as of solid achievement. This number contains an article sketching the developments shown in the JOURNAL's short history which should be of interest to all students of extra-sensory perception. Such a perspective of the field of work has brought out relations that may not have been apparent even to those who have read each issue of the JOURNAL as it has appeared since March 1, 1937.

"Why do you not use something else besides card tests?" the ESP experimenter is often asked. "Because the card test is simple, quick, and safe against error," is the usual answer. But now a test using drawings is presented by Stuart, which meets all three require-

ments. Especially since Stuart has shown this to be a successful method, we expect to find more use of drawings in ESP work in America in the future, supplementing the valuable work of Carington abroad.

A new measure useful in detecting ESP was introduced by J. B. Rhine in the September number of last year—the salience ratio, or SR. This measure is due to the discovery that the mere position of the trial in the run of twenty-five has, in a great deal of the ESP work, made a difference in the relative rate of success. The SR is based on this distribution of hits according to position. In the present issue it is used to evaluate the results of a transatlantic test of ESP, one involving the largest distance yet obtaining in ESP tests reported, to our knowledge. This work puts still greater burden on the question of the soundness of the method itself, and the empirical check on the method reported by Pratt, Humphrey, and Rhine in this issue will be highly pertinent to that question. This is needed all the more since other articles utilizing the SR method, even in tests of the precognition hypothesis, are scheduled to appear in the future.

* * * *

A final note may be added as to the long-run objectives of those whom this JOURNAL represents, as far as we can interpret them. From Pearl Harbor and the plains of Russia it is a far cry to the "Peace on Earth, Good Will toward Men" which still rings in our ears as these lines are written. But we now know that war is a form of imperfect social engineering, and that this is largely due to our ignorance of the "materials" to be engineered—our ignorance of man, his make-up, and his place in the natural order. The impulse to good will among men—our whole ethical structure—has been identified with faiths and doctrines which "endowed" man with those unusual capacities with which we are dealing in this research. And the moral guidance of the race has in the past derived from these beliefs, *as that of the future will depend on what can be discovered about our nature as human personalities*. Experimental confirmation of even a small part of the teachings of these earlier faiths justifies us in going ahead to determine whether the more traditional conceptions of human nature have still further support in experimental fact, or whether the apt analogies of modern physics apply to the world of value and experience, as their many advocates assert.

FIVE YEARS OF THE JOURNAL OF PARAPSYCHOLOGY

By DOROTHY H. POPE and J. G. PRATT

WITH THIS ISSUE, the JOURNAL OF PARAPSYCHOLOGY enters upon a new phase of its existence. Since the course which it is to take in the future will be conditioned in part by what it has been in the last five years, the present moment is an appropriate one for examining its development and noting whatever general relationships and directional trends have emerged during that time.

In undertaking such a survey, we are aware of the fact that some of the trends to be reviewed have been pointed out before, particularly in the editorial comment of the JOURNAL. The features of the past five years of ESP research which are to be noted herein have a more generalized character than the contents and findings of the individual articles, for they belong to the work as a whole rather than to the separate contributions. We feel, therefore, that even the closest students of this periodical will benefit by having these broader aspects of development as they are reflected in the JOURNAL OF PARAPSYCHOLOGY brought together in a single article.

A CENSUS OF CONTENTS AND CONTRIBUTORS

"I can't imagine how you can find material to fill such a journal, but time will tell." This was the comment of a prominent American psychologist in a letter to one of the editors of the newly launched JOURNAL OF PARAPSYCHOLOGY in 1937. Thus far, what *has* time told? In respect to quantity of material, the five volumes of the JOURNAL have held to the announced plan of having each number contain approximately 300 pages. And as for the future, there was never a time during the first five years when there was as much material awaiting publication in this periodical as there is today. The question of whether the quantity of material on experimental ESP

would be enough to support a journal has therefore clearly been answered in the affirmative.

But what about the more important question of kind and quality? The articles in Volumes I to V as they are listed in the Tables of Contents can be classified into groups along the following lines:

- Original research reports: 39
- Statistical articles: 14
- Articles dealing with criticism: 12
- Reviews of earlier research: 7
- Discussions of methods and apparatus: 5
- Symposia of psychologists discussing ESP: 2
- Opinion surveys directed to psychologists: 2
- Theoretical articles: 2
- Conditions favorable to ESP: 1
- Concerning hyperaesthesia, a related phenomenon: 1.

In addition, there are six research notes reporting original research, one note describing apparatus, and other discussions in the sections called *Notes* and *Letters to the Editors* which deal with pertinent topics of research methods, statistical procedures, and theoretical considerations.

It is evident that in accordance with the purpose editorially expressed at the outset, the JOURNAL has been predominantly a medium for reporting experimental findings. This is even more nearly the case than the articles themselves at first suggest, for almost all the contributions beyond the 39 original research reports are directly concerned with discussion and clarification of laboratory findings and the facilitation of further discoveries.

Purely speculative and theoretical discussion has been conspicuous by its absence. We believe it would be erroneous to take this fact to indicate that the investigators have been lacking in appreciation of the deeper significance of their findings. Rather has there been a studied effort to avoid undue theorizing in the early stages of the investigations. This effort has been made with the purpose of first collecting a sound and varied body of data upon which to base generalizations with correspondingly greater assurance. The two theoretical articles which did appear were written by authors who had not themselves contributed to the experimental results.

With few exceptions, the writers of the research reports concluded that their results constituted evidence of ESP. The research was conducted in eighteen different colleges or universities and five private laboratories. Thirty-five investigators have contributed original research reports to the JOURNAL and the work of four others has been reviewed in separate articles or notes. All but four of the 39 reached conclusions favorable to ESP on the basis of their own results.

The Tables of Contents indicate a tendency toward longer research reports, with fewer of them in each volume. The number of articles reporting or reviewing results in each of the five volumes has been in this order: 20, 12, 5, 3, 6. The greater number of reports in the first volume is explained in part by the fact that some of the experiments had been completed before the launching of the JOURNAL provided a medium of publication. While the number of research articles has declined, their length has increased so that the amount of material has remained about the same. This trend toward fewer, longer articles reflects the fact that the research has advanced to such a point of technicality that amateur or nonprofessional investigators have largely dropped by the wayside.

THE SEARCH FOR A "REPEATABLE" EXPERIMENT

One of the persistent tendencies in the research has been in the direction of discovering a "repeatable" experiment. (The quotation marks are used to convey some necessary cautions.) The aim of science is understanding and control, and we doubt the fruitfulness of searching for any kind of repeatability except that which would naturally come through the achievement of these ends. Of course, a repeatable ESP experiment might be discovered quite by accident, but its scientific value would remain strictly limited unless the recurrent success achieved thereby led to new knowledge of the *modus operandi* of the process. In this sense, every investigation in ESP is a search for a repeatable experiment. And the fact that the occurrence of the phenomenon has been confirmed so many times over gives an emphatically affirmative answer to the question: Is ESP repeatable?

Even in the most highly developed of the sciences, there are only degrees of repeatability. What some ESP investigators have meant,

therefore, when they spoke of discovering the repeatable experiment is that they aimed to conduct such tests and to make such analyses as would lead to a better understanding of ESP and a higher degree of control than had previously been achieved. No one has ever set invariable and unconditioned repeatability as his goal.

The recognition of this aim of greater repeatability as one of primary importance has become more explicit in the *JOURNAL* as it has grown. In particular, Stuart has been concerned to make tests and analyses with this purpose in mind and his studies have led to reports upon the relation of scoring to rate of response (II, 171)¹ and upon the relation of the subject's estimate of his scores and ESP performance (V, 99). Carington has also made explicit his aim to develop a repeatable experiment along the lines of his work with drawings (IV, 1). Murphy and Taves (III, 38) expressed the hope that other investigators would be able to duplicate the results which they obtained from subjects working on a variety of contemporaneously administered ESP tests.

Some of the critics have likewise urged repeatability as an aim of the research; their emphasis, however, has been upon whether the actual experimental procedures could be explicitly repeated by employing the information given in the reports.

DEVELOPMENTS IN TESTING METHODS

The story of ESP testing has been, in general, one of constantly increasing safeguards against all conceivable sources of error. The investigators reporting their work in the first volume of the *JOURNAL* went ahead under the assumption that the usual standards of precaution for scientific research would be high enough for use in the field of ESP. The later issues of the *JOURNAL*, however, reveal that the various safeguarding features of methodology have taken on larger (and sometimes inconvenient) proportions. That is not to say that progress in this direction has been perfectly regular. Some experiments, reported before the *JOURNAL* was started, were almost ideal on the point of safeguarding; while some that have been published at a later time have left something to be desired in this regard. In general, the emphasis upon formal precautions reached

¹ These figures refer to volume and page numbers of the *Journal of Parapsychology*.

its highest point in the Pratt-Woodruff report (III, 121), a level which has been maintained in practically all experiments conducted since that time. These safeguards were devised to eliminate various kinds of hypothetical errors in the experimentation. (The errors are called "hypothetical" advisedly, for their actual occurrence in the general ESP testing situation is something which has never been demonstrated.)

The first concern of the investigator in building up safeguards has been the elimination of sensory cues, such as faulty cards, whispering, and "peeking," as a possible basis of extra-chance scoring. Next has come an emphasis upon the avoidance of certain statistical pitfalls and of all sorts of clerical errors in recording, checking, counting, and computing. Finally, attention has been given to introducing controls against motivated errors on the part of individuals connected with the conduct of the experiments by making it difficult—if not impossible—for the results to be altered by one person without that fact being noticed by someone else.

The methods of eliminating sensory cues, as summarized by Rhine (I, 276), come under the headings: cards sealed in opaque envelopes, tests with screened cards, tests with distance and walls separating cards and subjects. These methods of control against sensory cues have been standard since that report, although in some later tests the question of cues as a factor in the subject's responses has been completely removed by the fact that the stimuli to be perceived were not selected until *after* the subject had made his calls.

More stringent safeguards of a statistical nature have been introduced mainly on two counts. One was to meet the criticism of unjustified selection of data. An experimenter might, by using only certain selected portions of his data, obtain spurious results in his experiment. In the first two volumes of the JOURNAL, the question of selection of data was covered by explicit statements from the investigators that all of the trials made were included in the reports. Later, articles from the Parapsychology Laboratory reported the practice of keeping all records on officially stamped, serially numbered sheets. By designating certain numbered record sheets in advance as intended for use in a particular experiment and by accounting for all of them at the time of reporting the results, the

investigators met the question of selection as objectively and with as much finality as was possible.

The second point was concerned with optional stopping, a practice by which an experimenter might, theoretically, lift his results to a level of significance by taking advantage of chance fluctuations in scores and stopping the test at a favorable point. This possibility was not considered in the earlier reports in the JOURNAL. When the point was raised by Leuba (II, 217), the question was taken up by statisticians and methods for making allowance for the effect were presented. The first application of a correction for optional stopping was made by Pratt and Woodruff (III, 121). Most investigators, however, have preferred to deal with the question by stating in advance the length of series to be conducted, as in the case of the Martin and Stribic reports (II, 23; II, 287), or by showing that the stopping-point could not have been affected by knowledge of the scores. In general, the publication of complete surveys with the pooling of all results for a given method took care of the problem in the period of publication preceding the launching of the JOURNAL, and the question is still effectively answered by occasional reports and surveys which summarize *all* the trials made under certain conditions. Such, for example, are the book, *Extra-Sensory Perception after Sixty Years*,² and the report by Rhine dealing with precognition (V, 1).

In the Pratt and Woodruff report, there appeared the first explicit controls against motivated errors. These comprised the joint checking of the scores by two experimenters at the time of the test, the practice of keeping duplicate personal records of all scores, and of putting the records under lock and key, beyond the reach of either investigator, until they were independently rechecked by a third person. No claim was made in the report that these controls were perfect, and doubt was expressed that perfect controls against motivated errors on the part of the investigators were practicable. In commenting upon this article, some members of the Board of Review of the JOURNAL expressed the opinion that considerations of this kind have no place in scientific reports. Nevertheless, the tendency to take explicit account of the question of both motivated and unmotivated errors has continued in some of the later articles,

² By Rhine, J. B., Pratt, J. G., Stuart, C. E., Smith, B. M., and Greenwood, J. A. New York: Henry Holt, 1940.

particularly those of Hutchinson (IV, 249), Rhine (V, 1), and Humphrey and Pratt (V, 267).

Perhaps the fullest degree of control in respect to safeguards would be reached with the perfection of reliable self-testing apparatus and automatic ESP machines into which the necessary precautions could be incorporated. The first number of the JOURNAL included a review by Smith (I, 63) of the experiments by Tyrrell which featured an automatic testing device. This apparatus was described in detail by Tyrrell in a later article (II, 107). The emphasis upon ESP machines reached a climax in Volume III which contained three articles on the subject. With the exception of the Tyrrell experiments, none of the efforts toward developing an apparatus for automatic selection of the stimulus and recording of the results have reached the stage of reported use in actual investigations, though the use of self-testing, semi-automatic devices has been reported by Humphrey and Pratt (V, 267).

THE TREND TOWARD MORE SENSITIVE MEASURES OF ESP

The most commonly used measure of ESP performance is the critical ratio, or CR, based upon the deviation of hits from the most likely chance number. By its use, the investigator arrives at a probability value which tells him how probable it is that success shown by the subject could have occurred purely by chance.

The critical ratio is the most obvious measure to use as long as it can be assumed that the subject's success is a matter of the exact and consistent hitting or missing of the stimuli toward which he is directing his ESP ability. There have been increasing evidences, however, that success in these tests sometimes takes more subtle forms to which measures of the total scores or averages are not sensitive—forms in which ESP ability may lie concealed in what appear to be chance scores.

One such subtlety of scoring is the tendency of subjects to show characteristic curves of hitting through the run. These curves have been given in several reports in the JOURNAL, and the results have generally confirmed the earlier findings of the Duke experiments that in calling "down through" a pack, more success is achieved at the beginning and end of the run than in the middle. If chance alone, and not ESP, were in operation, the subject would score as many hits

at the center as at the ends of the run. Rhine (V, 183) has given the tendency to score higher at the ends of the run the name of "terminal salience" and has, with the mathematical assistance of Greenwood (V, 245), used the manner of hitting throughout the run as evidence of the occurrence of ESP. It appeared from his results that the subject might actually have patterned his hits within the run by means of ESP even when the total number of hits in the experiment as a whole was not significantly different from what would be expected by chance.

In another experiment, Murphy and Taves (III, 38) noted that when subjects were assigned four different types of ESP tasks to perform in rotation during the same session, there was a tendency for the success in these tasks to vary up and down together to an extra-chance degree, though the score averages were not significantly high.

Still another instance in which ESP is reported to have occurred in an indirect and unexpected way is in Carington's experiment with drawings. He found that subjects tended to displace their hits, sometimes scoring successes upon drawings which had been used as targets on the immediately preceding days and sometimes hitting drawings yet to be selected for the following days. Such displacements appeared to be confined to the ten-day period covered by each test series. When the data were evaluated in these ten-day units, the results were found to be significant. This was not the case when each day's results were considered separately and the possibility of displacement had therefore been ruled out.

In the aforementioned work of Stuart, the search for a repeatable experiment has been associated with the use of more sensitive measures of ESP. He concluded that the rate at which the subject responds and the manner in which he estimates his scores are related to success in ESP tests.

This trend toward the use of new ways of detecting ESP is at its height and so far has shown no signs of abating.

THE TREND TOWARD TECHNICALITY

Any reader who compares the first volume of the *JOURNAL* with the last one cannot escape notice of the continual increase of technicality in the reporting of experimentation. In the first volume, the

experimental articles simply report straightforwardly the results obtained when subjects were tested for ESP after the manner of the earlier Duke studies. The efforts made to reassure the reader as to the authenticity of the conclusions in these first articles are such as might be found in any other psychological periodical. In the last volume, however, attention is given to technical points which almost make it appear that the articles are written with a totally new vocabulary. During the time represented by the intervening volumes, the investigator has become concerned about such questions as optional stopping, the stacking effect, the matching vs. the binomial hypothesis, covariation, displacement, and, in Volume V, salience. The reports give evidence of much greater technicality in testing methods and much more effort at enforcing, by explicit rules in the experimental procedure, those controls which had formerly been left to the competence and trustworthiness of the investigators.

This has naturally, if not necessarily, resulted in a narrowing of the circle of active investigators of ESP problems. The increased technicality of the field of parapsychology has made it more and more difficult to initiate new experimenters into the subject. To some extent, this growing complexity has been a direct result of the intense controversy over the research. But there may be a bright side even to this picture. If a trend toward technicality is one of the growing-pains of any pioneering research, the field of ESP, as reflected by the literature of the *JOURNAL*, has shown unmistakable signs of progress.

SCORE AVERAGES

One of the most consistent features of the research, according to reports in the *JOURNAL*, is a decline in the score averages reported by research workers. To what extent is this impression justified by the facts? What relation, if any, is there between loss of ability by individuals and decline of high scoring in the work as a whole? How have score averages been related to the adequacy of experimental conditions?

The odds-against-chance for that work reported from the Duke Laboratory before the *JOURNAL* was started were generally, and correctly, conceived to be astronomical. It seems safe to say, however, that most people had an exaggerated notion of the actual score averages upon which these astronomical odds were based. The highest-

scoring subjects mentioned in the book, *Extra-Sensory Perception*, published in 1934 by J. B. Rhine, made averages which ranged only from seven to ten per run of 25 cards when 5 was the number expected by chance. In exceptional shorter series, some subjects achieved averages in the low 'teens. The significance of the results lay in the fact that extra-chance averages of *even this order of magnitude* were maintained for series of great length.

The first confirmations of the Duke results which were published in the early numbers of the JOURNAL showed the same general pattern of score averages: subjects who had been selected on the basis of outstanding performance in preliminary tests were able to continue to score in the general range of the earlier results. Up to that time, therefore, there was no decline in scoring. In fact, the highest score average ever recorded is that of 18.24 hits per 25 trials for a series of 74 runs which was reported by Riess (I, 260) in the first volume. In the same number, Warner (I, 234) reported a series of 10 runs which averaged 9.3. Score averages had, if anything, risen.

The subsequent volumes of the JOURNAL do show fewer cases of an outstanding level of performance, though there is by no means a regular decline. The reports of Martin and Stribic include some of the best examples of sustained high scoring that have ever been reported. But there has been a notable tendency toward lower score average in contemporary reports. The chief reason for this is undoubtedly the shift of emphasis in the reports from individual performance by gifted subjects to the bearing of results as a whole upon a particular problem. With increasing frequency, authors show less concern with individual performance; so much so that in most of the recent articles in the JOURNAL it is impossible to see how the performances of the various subjects compare with each other. With less attention given to selecting persons of outstanding ability and with emphasis placed upon the total average for the experiment, it is natural that score averages should have declined.

It seems probable, however, that there are other factors involved as well, among them the increase in formality and technicality in the tests. Much of the spontaneous interest which was shown by subjects and experimenters alike in the earlier tests has been sacrificed by

explicitly restricting the subject's experimental freedom in order to furnish the increasing safeguards.

It is obviously incorrect to suppose, as one critic has done, that the scores have become lower because the safeguards have become more adequate. The work of Riess and of Warner involved both distance and independent recording, with the consequent elimination of sensory cues and recording errors. The work of Martin and Stribic as presented in their final, detailed report (IV, 159) has never been successfully criticized as failing to meet both these requirements.

In part, the failure to report higher scores in the most recent papers has been brought about by the investigators themselves. They have—whether for better or for worse—shown an increased willingness to use dragnet methods, relying more and more upon functions like displacement and salience to reveal the subtle or indirect forms in which ESP may find its expression.

CRITICISM

The critics of ESP research have generally preferred not to publish their articles in the *JOURNAL OF PARAPSYCHOLOGY*; in consequence, attention to criticism in the *JOURNAL* is chiefly in the form of surveys of attacks upon the work and its conclusions which have appeared elsewhere. At the end of Volume I, Greenwood and Stuart reviewed the mathematical criticisms of ESP as they stood up to that time, and at the ends of Volumes II and III Stuart had articles bringing the review of the criticisms up to date for the research in general. It was concluded in these reviews that none of the criticisms necessitated any change in any of the major conclusions of the research.

The year 1938 witnessed a high point of critical attention to ESP within the psychological profession. Mention is made in the "Editorial Comments" of the December number that ESP had been discussed at a number of sectional psychological meetings, and that was the year in which the A.P.A. held a symposium on the subject. There may be some dispute as to whether the profession was enthused or aroused, but of the fact that they were aware of ESP there can be no doubt.

The A.P.A. symposium was arranged and presided over by one of the most persistent critics of ESP. It quite naturally, therefore,

took on the form of a debate, with speakers *pro* and *con* on the questions of the mathematical basis of the results, of sensory cues, and of recording errors. The general topic for discussion was that of developing an adequate methodology for testing the ESP hypothesis. The upshot of the discussion was that the best methods already in use were fully adequate to determine the answer to the question of the occurrence of ESP.

Volume IV of the JOURNAL contains an exchange of controversial articles dealing largely with statistical aspects of ESP. Feller (IV, 271) raises the familiar criticisms of inadequate shuffling, selection of data, and favorable optional stopping of the experiments, but Greenwood and Stuart (IV, 299) show that the case for ESP could be based entirely upon work to which these criticisms are not applicable and that these questions could not account adequately for the results even in those tests in which theoretically they might have occurred. Three articles of minor criticism also appear in Volume IV, but there are no critical articles in Volume V. At the same time, there has been a noticeable falling-off of critical attacks upon the ESP work and conclusions in other journals. This, together with the fact that the more recent criticisms have largely been repetition of earlier ones of which the investigators had already taken full account, indicates a definite decline in adverse articles.

The Board of Review has, however, by virtue of its official capacity, made critical comments upon the original research articles of the last three volumes. The members of the Board found nothing to criticize in some articles which reached the conclusion that ESP occurs and were critical of others which reached the same conclusion. However, they were also critical of one article which attacked ESP, thus showing that they were consistent in applying their standards. The authors of papers which received adverse criticisms were sometimes able to make the additions to or the changes in their reports recommended by the Board; but in regard to many points raised, they were in turn critical of the critics and endeavored to show that the criticisms were not valid or relevant.

PROBLEMS AND THE EVIDENCE

In spite of the fact that the publication of evidence for and about ESP is the *raison d'être* of the JOURNAL, we shall not attempt to

summarize progress in this line in detail. This decision is made in the belief that most students of these pages will already have become informed beyond anything we could say in a limited space. To those who do not feel sufficiently acquainted with the results, we recommend the discussion of the evidence on the nature of ESP in Part III of *Extra-Sensory Perception After Sixty Years*. There are, however, certain aspects of the development of problem emphasis which come within the range of our review.

The primary emphasis of Volume I was upon the problem of the occurrence of ESP, particularly in regard to the number of independent confirmations of the positive findings of the earlier Duke experiments. Attention was chiefly given to testing for clairvoyance under a variety of conditions. In spite of the primary importance of the evidence from the point of view of proving that ESP is a real phenomenon, most of the experiments reported in the first year were such that, when significant results were obtained, they had a bearing upon problems concerned with the nature of ESP. Among the secondary questions which the evidence helped to answer were the following: How is ESP performance affected by differences in test procedures? What is the effect of differences in the kind of stimuli used? How do subjects of different ages and mental conditions compare? What is the effect of distance?

The other four volumes showed a continued emphasis upon the question of the case for ESP, but this problem was of less relative importance than before because of the force of the accumulated evidence. At the same time, the problems concerning the relations of ESP to other phenomena took on an increased significance. The problem of the effect of the experimenter-subject relationship upon scoring in the tests was investigated by Pratt and Price (II, 84) and by MacFarland (II, 160), and it was found that this social factor affects performance. Stuart's investigations, already mentioned in other connections, showed that the rate of responding may affect the subject's level of scoring (II, 171) and that subjects who estimate their next scores without regard for their past scores get better results than those whose estimates are influenced by the last scores made (V, 99). Other points of evidence as to the nature of ESP are that the level of performance may vary together in different tests con-

ducted in rotation (Murphy and Taves, III, 38), that novelty of the stimulus favors scoring (Pratt and Woodruff, III, 121), that subjects may unconsciously displace their ESP successes backward or forward in time to past or future stimuli in the same experimental series (Carington, IV, 1), that significantly low or negative scores may be obtained when the subjects are actually trying to get as many hits as possible (Rhine, V, 1; Humphrey and Pratt, V, 267), and that the hits may be distributed through the run in a significant manner even when the score averages are not themselves such as would be unexpected by chance (Rhine, V, 183).

A further feature of the evidence which deserves special mention—both because of its profound philosophical implications and because it represents such a strong emphasis in the research at present—is one which has grown out of the study of the time relations of ESP; namely, precognition. Two research reports dealing with this subject appeared in Volume II of the *JOURNAL* (Rhine, II, 38; Rhine, Smith, and Woodruff, II, 119), and there has been since that year an ever-increasing emphasis upon the problem of how the ESP process is related to time. Because of the difficulty of devising an adequate test of precognition, the successive articles are characterized by advances in methodology, particularly as regards the way in which the future stimuli to be perceived are selected. In the last two volumes there have appeared two original research reports devoted wholly to answering the question whether subjects can use ESP to predict future, unferrable, random events (Hutchinson, IV, 249; Rhine, V, 1), and there have been three other reports in which some evidence bearing on this problem has been offered (Carington, IV, 1; Stuart, V, 99; Humphrey and Pratt, V, 267). With the exception of the Humphrey and Pratt report, all of the investigations have yielded results which indicated the occurrence of precognition. The revolutionary nature of the problem, however, makes it one which will require further and possibly prolonged study for a final solution.

* * * *

A question which is raised forcefully by the backward glance at the first five years of the *JOURNAL OF PARAPSYCHOLOGY* is: What does it all amount to in the way of actual and permanent scientific accomplishment? What progress has been made toward getting the

field of parapsychology recognized as one of legitimate scientific inquiry, and to what extent have the conclusions reached in the research reports been accepted generally as forming a part of the established body of scientific knowledge?

The two surveys of professional opinion which were published in 1938 (II, 296; II, 302) showed that, while few psychologists among those queried were convinced that ESP is a real phenomenon, the vast majority of them thought that the problem of ESP is a legitimate one for scientific investigation. This professional judgment gives the ESP research worker reason to believe that his reports will in the course of time get a fair hearing. This fact, in itself, is of such importance that its significance for both the investigator and his field could scarcely be exaggerated.

The question of how widely the conclusions expressed in the *JOURNAL* have been accepted is one which cannot yet be answered. Of course, the greatest sustaining hope of the scientific explorer is that his work will ultimately be accepted as having general validity, since otherwise whatever deeper significance it might have could never be fully realized. Without claiming to find definite, measurable evidences that any particular percentage of the readers of the *JOURNAL* are convinced of the occurrence of ESP, we find a certain significance in the fact that so many investigators have reached this conclusion in the reports of their own experiments. Some may have undertaken their own investigations with a leaning toward acceptance of ESP on the strength of earlier evidence, but these were probably in the minority among the 35 experimenters who reached conclusions favorable to the occurrence of ESP in this periodical alone. When we consider that perhaps at least as many more have published results of similar bearing elsewhere and that the investigators, mostly psychologists, have consisted largely of young people whose full influence among their professional groups is yet to be felt, the prospects for the wider influence to be secured from the better-established findings of parapsychology are distinctly encouraging.

AN ESP TEST WITH DRAWINGS

By CHARLES E. STUART

ABSTRACT: The test of ESP in which a percipient draws or describes a randomly selected object looked at by an agent is an old and well-tried method. The statistical evaluation of the difference between a successful performance and the occurrence of accidental similarities has been a more recent problem. This paper reports a method for such evaluation by a matching technique in which the agent and percipient act separately as the judges of the objective degree of correspondence which exists between the stimulus drawing and the percipient's response. The judgments involve matching each stimulus drawing separately to the response drawing, and giving the latter a rank order of choice as to which is the best match, which second, and so on.

In the research reported here, the method is applied to seven experimental sessions, each involving four drawing responses by a subject. The cooperation of a purported "good ESP subject" was secured as percipient in two of the experiments and as agent in the others. Various members of the laboratory staff served as the other subjects.

Without his knowing the order of the stimulus drawings the percipient was instructed to match them preferentially as stated above. Then the agent matched the drawings preferentially without knowing the order of the responses. The number of true correspondences denoted as first and second choices was over twice the number of true correspondences evaluated by the judges as third or fourth choices. A score was calculated which was amenable to statistical evaluation. Evaluation with respect to each of three chance hypotheses yielded statistically significant difference of the experimental results from chance expectation. The "good subject's" judgments were independently significant.

The suggestion is offered that the conception of the free drawing method as a means of recording a percipient's response to a stimulus situation, instead of the limited conception of his drawing as an attempt to reproduce the stimulus drawing, will lead to a greater degree of freedom for the percipient in his response. The unrestricted mode of response may possibly have accounted for the fruitful results of this very brief investigation.

THE ESP test situation in which the percipient tries to reproduce a figure or drawing which is looked at by an agent or sender is one of the classical methods in psychic research. It presents objectively and in very simple form the kind of content one would ordinarily expect would be transferred in a case of "thought-transference." Early studies utilizing this method are very convincing in the accuracy of the subjects' drawings as they reproduce the originals. Some of the most interesting of these are found in Upton Sinclair's *Mental Radio* (4), an important discussion of which is offered by Walter Franklin Prince in a bulletin entitled "The Sinclair Experiments" (3). (The latter source also includes a short history of drawing experiments in general.) Prince and others developed a method of evaluating the accuracy of correspondence between original and reproduction, a method based upon comparisons of the work of the original subject with that of several control subjects supposedly lacking in ESP ability. Some time later, Pratt, reporting some observations by C. Hilton Rice, proposed the method of correct matchings as a means of evaluating free drawings (2), in which an independent judge, not knowing the true correspondences, matched original drawings and subjects' reproductions together in the way they seemed best to correspond. The method of statistical evaluation had later to be revised because of its lack of validity when more than one judge was employed. Recently, the English investigator, Whately Carington, conducted a large series of free drawing experiments, in which some of the more successful results involved percipients in America attempting to reproduce the stimulus drawing in England (1). These were evaluated by having an independent judge assign the originals and the reproductions two degrees of cor-

respondence (a hit or a half-hit). The judge in this case did not know which drawings were intended for which originals.

The evaluation of subjects' response drawings presents certain theoretical problems which, if we are to get efficient results from the statistical measures, must be considered in terms of assumptions regarding the nature of the transferring process; that is, the nature of ESP. If we compare the responses of a subject, who by reputation or accomplishment in some form of ESP test is expected to yield consistently successful results, with those of subjects without such reputation, we make the natural assumption that certain people are more "gifted," and consistently so, than others, and that we are able to select them with some reliability. With such an unusually successful series of experiments as those of the Sinclair case, the Prince procedure of evaluation on this assumption might be a very valid method. The Carington experiments, on the other hand, utilized many unselected subjects, and the problem of evaluation there lay in finding whether they generally made more hits than would be expected by chance. The evaluation of a given reproduction involved, therefore, its classification as a hit or a miss (or a third classification, a partial hit). This evaluation assumes that the evidential features of the drawing which are transmitted or perceived should be sufficiently apparent and clear-cut that they can be given an all-or-none classification. Since this is not a necessary assumption regarding ESP, it is apparent that the method suffers from just that degree of crudity of evaluation.¹

The method of correct matching requires no evaluation of the degree of correspondence. Its sensitivity is greater than classification methods in that a judge has placed before him a given number of originals and a given number of reproductions which must be matched, one to the other respectively, until all are paired up. He may use, if necessary, small similarities of any kind to justify the matching. But again, in this procedure, the relation of a given drawing to a

¹ These remarks on Carington's method are concerned with the simple hypothetical aspects noted, and are not to be construed as "criticism." Actually Carington utilized three different methods of evaluation, one of which was the method of correct matching. He found the "method of palpable hits" the most effective in dealing with his data. Any thorough review of drawing evaluation would inevitably require extensive reference to Carington's excellent discussion of "Methods of Assessment" (1).

given original is treated in evaluation as an all-or-none item; that is, if a drawing, "A," is matched with original, "X," statistically this matching is either a hit or a miss, and will be so tallied.

However, a practical difficulty arises almost immediately when the method of correct matchings is used. If the original drawings are sufficiently discrete items, and the subjects' reproductions are clearly differentiated from one another and match the originals to a striking degree, the method gives good results. But, as will occur if drawings are randomly selected, frequently the originals may resemble one another to a great extent. Likewise, the subject commonly will draw in any experimental period two or more objects which are very similar in form or meaning. Suppose one of these is the correct reproduction of the original, the judge may match it correctly fifty per cent of the time, but he will also match the unappropriate but similar drawing fifty per cent of the time. Since the statistical treatment requires the hits to be considered in an all-or-none fashion, a quite striking hit might be devaluated by a purely chance occurrence in the selection of the originals or by a fortuitous tendency toward repetition by the subject. From another point of view this may be considered a virtue of the method since it prevents over-evaluation of accidental hits of the same kind. But, if we are to take seriously the possibility that ESP may occur in a sporadic fashion in any given series, this virtue lowers the sensitivity of the method.

The method utilized in the work reported here was attained by combining the best features of the matching method with the best features of the method in which the impartial judge attributes a degree of correspondence between original and reproduction. In this combination matching procedure, the judge matched each of the four drawings of a set against the four originals of each experiment, and thus rated the degree of goodness of matching at four levels—the original, which matched best, second best, third best, and poorest. In this case, suppose that by chance a bird had occurred twice among the originals, and suppose the subject had drawn a bird. He would still find it necessary to choose which one of the originals was the best, but in doing so, and by giving a second best choice to the other, he would be giving some statistical weight to the fact that the second

choice was a better match than the other two. In the same way, the same judge matched each original against all four reproductions. The statistical evaluation of the results of such a matching against the chance hypothesis is straightforward, and will be discussed later.

A GENERAL ACCOUNT OF THE EXPERIMENT

An excellent opportunity to test out this method was offered recently at the laboratory. Mr. William Russell, an assistant in the laboratory, had made some ESP studies of a seventeen-year-old freshman college student, PH. PH had evidenced in a number of informal cases a rather consistent ESP ability. Mr. Russell made it possible for PH to visit the laboratory for a few days, and during that time we had an opportunity to try out free drawing experiments with this supposedly good subject. PH proved to be a very cooperative subject, and seemed keenly interested in what was to him a new experimental method. In carrying out the judgments, which can become very tiring work, he showed an extraordinary degree of patience and persistence.

During his visit, I was able to use PH in seven experimental sessions, which will be denoted hereinafter by the letters, A, B, C, etc. The procedure for Experiment A was as follows: I prepared four drawings, a face, a leaf, a map, and a child's block.² Russell, acting as agent, was seated in the room adjacent to the room where the percipient was to do his drawing. I gave the four drawings already shuffled to Russell, with the instructions to shuffle them further and to utilize them in that final order. He was to look at one until I called to him to look at the next one. (There is a double Masonite partition between the two rooms, with no doors, and one has to call loudly to communicate intelligibly.) The percipient, PH, was given several sheets of letter-size paper and instructed to draw whatever came to mind in regard to the drawing being looked at by the agent in the next room. Each drawing was laid aside after it was completed and the next one commenced after I had shouted to the agent, "The next one!"

² These were not "random" selections, but systematically chosen as representing four distinctive "ideas." In the later experiments the drawings were more randomly selected from an unordered stock of 600 drawings. The method of evaluation does not require actual random selection of stimulus objects, but does require that the *order* of the stimulus objects be unknown to the percipient.

The percipient's drawings were apparently quite unrelated to the originals the agent was looking at. The first was a tree beside a hill; the second, an apple; the third, the outline of a house; the fourth, a heart. After these were completed, I went into the next room and received the shuffled originals from the agent, Russell. He remained in that room while I returned to the percipient. The percipient's four drawings were then placed before him and one of the originals given him. He was instructed to say which one of his drawings this original matched the best; which the second best, which the third best, and which the poorest. I marked his judgments as "1," "2," "3," and "4" on a prepared chart. After he had made these sixteen judgments, the four originals were placed face up before him and he was asked to match each of his drawings, one at a time, against the four originals with the same method of scoring. After this, he retired to a corner of the room, and the agent, Russell, was brought in and asked to carry out the same matching procedure.³ After he had concluded the matching procedure, the actual order of the originals as he had originally looked at them was compared with my record of the order of the reproductions. It was noted that the percipient, PH, had matched the original leaf as first choice for his tree and mountain. He had matched the original face as first choice for his apple. On the converse presentation he had also matched these pairs as first choices. These were correct correspondences.

Of the total of thirty-two correspondences presented to him, PH had evaluated the eight correct correspondences with four first choices, three second choices, and one fourth choice. The total score for the matching was 14, for which the chance expectation on a chance hypothesis is 20, indicating that he matched the drawings at a rate somewhat, but in no way crucially, better than chance.

In a second experiment, Experiment B, PH acted as agent and Russell as subject. The procedure was otherwise identical, except that when the agent made his judgments the subject left the room completely. The results in Experiment B were the best of any ob-

³ There was, of course, no communication between PH and Russell until after the experiment. The obvious step of guaranteeing the precaution by having PH entirely out of the room was not imposed here, because I had planned to use only PH's matchings as the final experimental data. Only after Russell had started did it become clear to me that judgments by the other participants would be equally as valuable as those of the subject being tested.

served. Russell, as subject, indicated five first choices for correct correspondences and three second choices. PH, the agent, making the same judgments, gave seven first choices and one second choice for the correct correspondence. It is notable in this experiment that neither judge designated a correct correspondence as a third or fourth choice.

The percipient's responses in this case consisted of numerous quick sketches of freely associated items. The subject ceased drawing and writing only when he had filled most of the page with various ideas.

In Experiment C, PH acted as agent and I was subject. In this case, the drawings were selected at random by Dr. J. G. Pratt from a box of 600 unordered drawings. My part in the procedure was then taken by Russell. Otherwise, the procedure was as before. In this case, my judgments gave the correct correspondences four first choices, three second choices, and one fourth choice. PH's matching of the same drawing yielded four first choices, three second choices, and one third choice.

These were the best of the seven experiments in terms of results. A total summary is presented in Table 1. Experiment D was performed with Dorothy H. Pope, the laboratory secretary, as percipient; Experiment E with Wallace Scherer, a research assistant, as subject. In Experiment F, I acted as agent and PH as subject. In Experiment G, PH was agent and Russell was subject.

Experiments A and B were conducted on the first morning of experimentation. Experiment C was done the same afternoon. Experiments D and E were done the following day in morning and afternoon respectively. Experiments F and G were carried out in the morning of the third experimental day. The time necessary for one complete experiment varied from thirty to ninety minutes, depending upon the complexity of the drawings by the percipient and the time taken for the judgments.

The foregoing gives a summary sketch of the investigation. But the various aspects deserve some fuller discussion before the results are considered.

THE STIMULUS SITUATION

The agent was seated at a desk in a room and was given the stimulus drawings on four cards enclosed in a folded paper. He was instructed to shuffle the cards and to look at the first one until told to look at the next, and so on. He was to record the order on the paper in which they were enclosed, and to retain this order after the conclusion of the period. He was to shuffle the cards well before returning them to the experimenter. The agent was alone in the room throughout the experiment, except for the two visits by the experimenter to give him and later receive from him the stimulus cards.

In concentrating upon each card the agent was told to keep the figure clearly in mind and in his field of vision. He was free to use any technique he chose for keeping his attention on the card, such as looking at various parts of it, imagining tracing it, etc. But he was not to "strain" in any way to achieve concentration. It was even all right if his thoughts wandered occasionally.

THE RESPONSE SITUATION

The customary instruction to the subject in experiments of this kind is that he try to draw what the agent is looking at. But it is evident that with the method used here this instruction is unnecessarily restrictive. If we conceive the subject's drawing as a *response* to the stimulus situation, there are many ways of responding equally as valid as reproduction of the stimulus. It is even reasonable to suppose that there are many responses much better than a reproduction, since in every-day life purely imitative behavior constitutes a very small fraction of our total responses.⁴

The instruction was, therefore, to "draw anything you think of. Don't worry about your skill in drawing. Just put down whatever comes into your mind in regard to the drawing the agent is looking at."






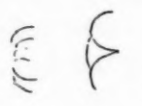

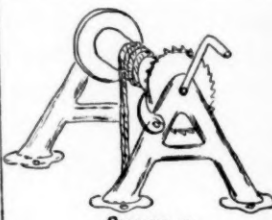
These instructions were interpreted in many different ways by the various percipients. One, WS, persisted in drawing, with considerable skill, a single object in great detail. Another, WR, recorded a large number of objects in a free-association style. The cost of

⁴ It is strange that in the many reports of drawing experiments that have been done, I know of no explicit statement of this obvious point.

reproducing all the results here is prohibitive, but Plate I gives several examples, and Table I presents a descriptive account of the responses.

Plate I

EXAMPLES OF TYPICAL STIMULUS AND RESPONSE ITEMS

| | | | |
|---|--|---|--|
|  <i>Exp. A, Stim. B</i> |  <i>Response b</i> |  <i>Exp. B, Stim. A</i> |  <i>Response a</i> |
|  <i>Exp. C, Stim. B</i> |  <i>Response b</i> |  <i>Exp. E, Stim. A</i> |  <i>Response a</i> |

There is a natural temptation to the experimenter to point out to the percipient who restricts himself to one object that he does not need to do so; and to suggest to the overproductive percipient that his proliferation of objects is going to make the matching task much more difficult than if he would record only salient impressions. It might be reasonably proposed that more complete instructions would yield more uniform kinds of responses and make them thereby more comparable one to the other. But again, if we conceive the percipient's drawing as a complex response it is evident that one subject may differ from another in many critical ways; and so long as we do not know the best mode of response it is wiser to leave the percipient complete freedom to develop his own.

THE MATCHING PHASE

If the conception of the percipient's drawing as a general response rather than an attempted reproduction of an object in the agent's

possession can be considered a basic feature of the method used in this study, an even more important second feature is the notion that the best judges of the degree to which a given response matches a given stimulus are the two people most intimately involved; namely, the agent and percipient. Again, this is a commonplace assumption in everyday life. A person knows his own experience better than anyone getting a second-hand account of it.

Instead, then, of presenting the stimulus and response drawings

Table I

DESCRIPTION OF STIMULUS AND RESPONSE DRAWINGS

| | Stimulus Drawing | Response | Comments |
|---|---|--|--|
| Exp. A WR Agent PH Percipient | A. Oak leaf. | a. Hill and tree. Wavy lines suggesting stream. | Contour of apple closely approximates contour of face. |
| | B. Man's face with right eye missing. | b. Apple. | |
| | C. Map of New York State. | c. Roughly sketched house. | |
| | D. Child's block with letter "A" on it. | d. Heart. | |
| Exp. B PH Agent WR Percipient | A. Psychological diagram. | a. "Pigeon," profile, looped thread, wavy lines, plus sign, "cucumber," "ace," "tiger." | A plus sign appears in both A and a. |
| | B. Key. | b. Telephone pole and wires, lunch box, "derby hat," "pheasant," "screen door." | Predominant food items in responses. |
| | C. Wineglass. | c. Lunch box, "cupboard," "Mother Hubbard," "dish," "banana," "walnut," "finger," brackets and miscellaneous marks. | |
| | D. Flag with crescent moon and four-pointed star on it. | d. Sun with corona, clock with hands indicating 9 o'clock, figure diving off board into pool, "Beach at Bali Bali or Waikiki," map of California, "rug," "ESP," roughly oval figure. | |

Table 1 (Continued)

| | Stimulus Drawing | Response | Comments |
|---|---|---|---|
| Exp. C PH Agent CS Percipient | A. Racquet and tennis ball. | a. Face of man with moustache, roughly drawn dog, Andy Gump face. | "Moustache" theme similar to dark lines across throat of racquet. |
| | B. Anchor. | b. Three sets of curved lines, one marked "umbrella," "ant or roach." | Form of anchor carried out closely. |
| | C. House. | c. Beacon light, "lighthouse," closed hand with thumb pointing right. | Fingers of hand show similar form to smoke coming from chimney of house. |
| | D. Sailboat. | d. Cross set at angle, cross drawn horizontally, outline of Lacrosse bat, "grass rake." | |
| Exp. D PH Agent DP Percipient | A. Telephone. | a. "Star-fish," burning candle in candlestick. | Some form similarity between candlestick and phone. |
| | B. Devil. | b. Ferns in flower pot, "fern-(ia)," rolling pin, chart of small circles. | |
| | C. Pointer consisting of stick lying across crooked branch. | c. "Shooting box, makes a hollow sound," a bow. | Both judges thought pointer was a sling-shot to shoot arrows. |
| | D. Feather. | d. Hat with feather, flag with musical symbols. | |
| Exp. E PH Agent WS Percipient | A. Hammock between two trees. | a. Windlass. | All responses skillfully drawn in careful detail. Form similarities apparent, but are well hidden in the structure of the drawings. |
| | B. Human figure made up of X's. | b. Statue of Greek athlete. | |
| | C. Fenced square like boxing ring. | c. Sail boat. | |
| | D. Oval. | d. Peach. | |

Table 1 (Continued)

| | Stimulus Drawing | Response | Comments |
|---|------------------------------|---|---|
| Exp. F CS Agent PH Perceptant | A. Andiron. | a. "Baby," "tree," moon, "cat," X, V, wavy lines, circle, "church," "alligator." | |
| | B. Safety pin. | b. "Ape," oval, barrel, girl's face, "penny," "ear," "8," curved lines. | |
| | C. Psychological diagram. | c. "Town," smoke stack, semi-circle, "match," "blanket." | |
| | D. Bunch of grapes. | d. "Camel," flask, star, thumb-tack, box, girl's face with "turban," "pertaining to Far East," hanging drapes. | |
| Exp. G PH Agent WR Perceptant | A. "No Fishing" sign. | a. "Saw," "apparatus," triangle, "tortoise," two graves with grave-stones, four lines, "bushel," "well-bucket too," "portrait," "slot machine," "loose-leaf" pad, "hoe," "moccasin snake or shoe." | |
| | B. Tic-tac-toe chart. | b. "17," "rose," calendar with automobile and January dates, "mosquito," triangle, "leopard," "Argentina," "ferry-boat," "ribbon," "lines," "glass," "napkin," "Christian Endeavor pack of hoodlums." | Part of chart apparent in drawing of automobile. |
| | C. Button. | c. Scissors, "cut," "Life in N. Y. Sept.," "bed," with reclining figure, ladder, "dog," Togo, three concentric circles, "tobacco," "maiden name," "144, 359," "jail," "pomegranate, kind of fruit," "40." | Both judges thought concentric circles were important item of similarity. |
| | D. Roast chicken on platter. | d. Square figure, "moon," rising over the hills, "watermelon," "honey," "cake," "potato," "tablet," needle and thread, "Santa Claus," "diamond," "jet black," "j." | Hills close in form to drawing of chicken. Marked food reference. |

to an independent judge to be evaluated as to degree of correspondence, the present investigation was arranged to permit the presentation of the drawings for evaluation to the agent and percipient themselves as judges.

By keeping the order of the stimulus drawings from the percipient's knowledge, he is able, as a judge, to match each on the basis of actual similarities or any other principle of correspondence.⁵ The main problem for the judge is to determine the principles of correspondence to follow in his selection. Here the instruction was simply "to decide which of your drawings matches or belongs with this stimulus drawing the best. Then which matches the next best, and so on." The subject was free to choose similarity of form, meaning, part structure, causal relation, even in some cases to make an intuitive guess without objective justification, to determine his judgments. The only restriction insisted upon was that each judgment be made on its own merits without respect to previous judgments.

There is the temptation to the experimenter, who is apt to be more familiar with the types of similarity to look for than the subject, to interpose suggestions to help the subject decide. Since the experimenter himself does not know the true order, this is not a violation of precautionary conditions. But it might violate the principle that the participants are the best judges.

My own lesson in this was the following instance: In Experiment B with PH as agent and WR as percipient, one of the stimulus drawings was that of a button, somewhat squared on the sides instead of circular, and having two holes in the center. In three of WR's drawings some similar form was evident, and I was surprised to find that he designated after much hesitation the other drawing as first choice on the grounds that three rather large concentric circles seemed the most similar item. After the matching I called the neglected form to his attention (insisting, of course, that I had no desire to alter his judgment ratings). He confessed he had not noticed the indicated similarity, and asked if he might change the order of one of his judgments. I agreed that he might do so. When PH was confronted with the same set of choices, he immediately matched

⁵ It is obviously possible, assuming ESP, that the judgment itself may be influenced by extra-sensory impressions. This likelihood cannot be ruled out, any more than it can from customary sensory perception.

the button with the concentric circles, with the remark, "That was just the idea I wanted him to get!" It developed later that this was the correct correspondence.

As noted previously, when the percipient acted as judge, his four response drawings were laid out before him, and he was given the stimulus drawings one at a time. For each of the stimulus cards, he had to designate which one of his response drawings was the best match, which second best, which third best, and which poorest.

Then the four stimulus cards were placed before him, and he was given his response drawings one at a time to decide which response was the first, second, third, and poorest match for each.

In the presentation of the same matching task to the agent, care was necessary to randomize the percipient's drawings to prevent influence of this order. This was accomplished by the use of simple unordered codes, the key to which the experimenter kept in his pocket until the final check-up.

RESULTS

In each of seven experiments four drawings were made. Each of these experiments involved two sets of judgments made by each of two judges. The actual observations are arrayed in Table 2. During the experiments the judgments were recorded in a random fashion, but they are displayed here so that the letters represent the serial order of the stimulus drawings and responses.

If we refer to the first block of sixteen judgments, the meaning of the table will become clear. The percipient had his four response drawings before him: *a*, *b*, *c*, and *d*. He was handed stimulus card C and asked to designate which of his drawings it matched best. He decided upon *a*, so the experimenter scored that correspondence as "1." He denoted *c* as second best, *d* as third best, and *b* as poorest. A "2," "3," and "4" were recorded in the appropriate squares.

The true correspondences are *A-a*, *B-b*, *C-c*, and *D-d*. In the instance noted they are scored as two first choices, a second choice, and a fourth choice respectively, with a resulting "score" of 8.

The scores of the seven experiments are displayed in Table 3. Since one of the purposes of the investigation was to test the purported abilities of PH, his scores have been listed together irrespective of his role of agent or percipient. If there were no observable

Table 2

RESULTS OF PREFERENTIAL MATCHING BY PARTICIPANTS AS JUDGES

| EXP. A | | EXP. B | | | | EXP. C | | | | | | | | | |
|-----------|---|-----------|---|-----------|---|-----------|---|-----------|---|-----------|---|---|---|---|---|
| PH, perc. | | WR, agent | | PH, agent | | WR, perc. | | PH, agent | | CS, perc. | | | | | |
| A B C D | | A B C D | | A B C D | | A B C D | | A B C D | | A B C D | | | | | |
| S to R's | a | 1 | 3 | 1 | 3 | a | 3 | 3 | 2 | 4 | a | 2 | 4 | 1 | 4 |
| | b | 2 | 1 | 4 | 2 | b | 4 | 1 | 3 | 2 | b | 4 | 1 | 2 | 3 |
| | c | 4 | 4 | 2 | 1 | c | 4 | 4 | 3 | 1 | c | 3 | 2 | 3 | 2 |
| | d | 3 | 2 | 3 | 4 | d | 2 | 1 | 1 | 3 | d | 1 | 3 | 1 | 1 |
| R to S's | a | 1 | 4 | 2 | 3 | a | 2 | 4 | 1 | 3 | a | 1 | 2 | 3 | 4 |
| | b | 3 | 1 | 4 | 2 | b | 1 | 2 | 4 | 3 | b | 4 | 1 | 3 | 2 |
| | c | 3 | 4 | 2 | 1 | c | 4 | 3 | 2 | 1 | c | 1 | 3 | 2 | 4 |
| | d | 3 | 1 | 4 | 2 | d | 2 | 1 | 3 | 4 | d | 2 | 1 | 3 | 4 |
| EXP. D | | EXP. E | | | | EXP. F | | | | | | | | | |
| PH, agent | | DP, perc. | | PH, agent | | WS, perc. | | PH, perc. | | CS, agent | | | | | |
| A B C D | | A B C D | | A B C D | | A B C D | | A B C D | | A B C D | | | | | |
| S to R's | a | 4 | 4 | 3 | 4 | a | 4 | 1 | 2 | 4 | a | 4 | 4 | 4 | 1 |
| | b | 2 | 2 | 4 | 3 | b | 4 | 1 | 1 | 3 | b | 4 | 4 | 1 | 1 |
| | c | 3 | 1 | 2 | 2 | c | 1 | 4 | 3 | 3 | c | 2 | 3 | 2 | 4 |
| | d | 1 | 3 | 1 | 1 | d | 3 | 3 | 4 | 2 | d | 1 | 1 | 3 | 2 |
| R to S's | a | 1 | 2 | 3 | 4 | a | 4 | 3 | 2 | 1 | a | 4 | 1 | 3 | 2 |
| | b | 2 | 3 | 4 | 1 | b | 4 | 1 | 2 | 3 | b | 4 | 3 | 2 | 1 |
| | c | 4 | 2 | 3 | 1 | c | 1 | 4 | 2 | 3 | c | 1 | 2 | 3 | 4 |
| | d | 3 | 4 | 2 | 1 | d | 3 | 2 | 4 | 1 | d | 1 | 2 | 4 | 3 |
| EXP. G | | | | | | | | | | | | | | | |
| PH, agent | | WR, perc. | | | | | | | | | | | | | |
| A B C D | | A B C D | | | | | | | | | | | | | |
| S to R's | a | 3 | 4 | 2 | 4 | a | 2 | 4 | 3 | 4 | | | | | |
| | b | 1 | 2 | 4 | 2 | b | 1 | 3 | 2 | 3 | | | | | |
| | c | 4 | 1 | 1 | 3 | c | 4 | 1 | 1 | 2 | | | | | |
| | d | 2 | 3 | 3 | 1 | d | 3 | 2 | 4 | 1 | | | | | |
| R to S's | a | 2 | 4 | 1 | 3 | a | 1 | 4 | 2 | 3 | | | | | |
| | b | 1 | 2 | 4 | 3 | b | 1 | 3 | 2 | 4 | | | | | |
| | c | 4 | 3 | 1 | 2 | c | 2 | 1 | 3 | 4 | | | | | |
| | d | 1 | 4 | 3 | 2 | d | 3 | 4 | 2 | 1 | | | | | |

Stimulus cards are designated in their true order by capital letters A, B, C, and D. Response drawings are designated in their true order by small letters a, b, c, and d. The preferential matching is indicated by the numbers. 1 indicates the subject thought the denoted S and R belonged together; 2 indicates his second choice; 3 his next choice; and 4 his designation of poorest matching.

Each experiment was matched four times, twice by each participant. The percipient matched each S to the four R's, then matched each R to the four S's. The agent did the same.

The score of each matching is the sum of the diagonals, Aa, Bb, Cc, and Dd.

relations between the stimulus cards and the response drawings, the scores should average around 10 as a mean. In the case of PH's scores at least, which average 7.6, they quite evidently do not approximate this chance average.

It is further evident from Table 2 that the true correspondences

Table 3

STATISTICAL EVALUATION OF THE TRUE CORRESPONDENCES

| Experiment | SCORES WITH PH AS JUDGE | | SCORES WITH OTHERS JUDGING | | Total PH | Total Others | Total All |
|---|-------------------------------|-------------------------------|-------------------------------|-------------------------------|-------------|-----------------|--------------|
| | 1st Matching S's to R's | 2nd Matching R's to S's | 1st Matching S's to R's | 2nd Matching R's to S's | | | |
| A..... | 8 | 6 | 11 | 10 | | | 35 |
| B..... | 5 | 4 | 6 | 5 | | | 20 |
| C..... | 7 | 6 | 7 | 7 | | | 27 |
| D..... | 9 | 8 | 11 | 11 | | | 39 |
| E..... | 8 | 8 | 12 | 10 | | | 38 |
| F..... | 11 | 13 | 10 | 11 | | | 45 |
| G..... | 7 | 7 | 7 | 8 | | | 29 |
| Total dev. from average of 10..... | -15 | -18 | -6 | -8 | -33 | -14 | -47 |
| Assump. of Independence CR. | 2.54 | 3.04 | 1.01 | 1.35 | 3.94 | 1.67 | 3.97 |
| Assump. of Dependence CR. . | 2.31 | 2.78 | .93 | 1.23 | 3.60 | 1.53 | 3.63 |
| Hypoth. of Correlation and Dependence CR. . | | | | | 2.84 | 1.23 | 2.42 |

are matched as first and second choices many more times than as third and fourth choices. Table 4 summarizes this observation, showing that the true correspondences were first and second choices more than twice as often as they were third and fourth choices.

Table 4

DISTRIBUTION OF JUDGMENT SCORES FOR TRUE CORRESPONDENCES

| Rank of Choice | PH AS JUDGE | | OTHERS AS JUDGES | | Total |
|-------------------|----------------------------|----------------------------|----------------------------|----------------------------|-------|
| | 1st Matching S's to R's | 2nd Matching R's to S's | 1st Matching S's to R's | 2nd Matching R's to S's | |
| 1..... | 11 | 13 | 9 | 10 | 43 |
| 2..... | 10 | 7 | 8 | 8 | 33 |
| 3..... | 4 | 6 | 5 | 4 | 19 |
| 4..... | 3 | 2 | 6 | 6 | 17 |
| 1 and 2..... | 21 | 20 | 17 | 18 | 76* |
| 3 and 4..... | 7 | 8 | 11 | 10 | 36* |

*A chi-square evaluation of the predominance of 1st and 2nd choices over 3rd and 4th choices gives a p-value of approximately .0001, which indicates a very significant departure from chance expectation.

STATISTICAL EVALUATION

Since every judgment consists of a designation of number 1, 2, 3, or 4, it is clear that the chance probability of a given judgment having one of these numbers associated with it is $\frac{1}{4}$. Also since a given score consists of the sum of four of these numbers, a sum which might be anything from 4 to 16, the mean chance expectation of the sum of four judgments is evidently 10. If every judgment is an independent item, the theoretical variance of the scores is 5. The standard deviation is thereby $\sqrt{5}$ or 2.236. The distribution is symmetrical, and rapidly approaches normality as scores are added together.

Assumption of Independence

Upon the hypothesis that every matching is independent, the sum of the deviations from a mean of 10 may be divided by the standard deviation of $\sqrt{5N}$ (wherein N is the number of scores) to get a critical ratio. These CR's are given in Table 3. Both of PH's sets of judgments have CR's significantly above a chance criterion of 2.50. When all the first judgments are totaled and all the second judgments totaled, each gives a CR greater than 2.5. And the total of all judgments gives a CR of 3.97, which is highly significant.

Assumption of Dependence

The judge is instructed to make each matching independent of all others. A given drawing may match several originals the best of the four. If the judge carries out the instruction exactly the assumptions of the above evaluation procedure are met completely. But we must consider the possibility that the instructions might be intentionally or unwittingly violated, and the effect such deviation should have upon statistical interpretation.

An extreme case exemplifies the problem. Suppose a subject gives first choice to the correct correspondence for each of four drawings. If each of these judgments were independent, the chance probability of such an accurate matching would be only one in 256. But suppose the judge had reasoned that, since he had given first choice to drawing *a* as matching original *A*, he should avoid giving first choice to drawing *b* against original *A*. And both *A* and *B* are

excluded from first consideration when he is matching drawing c. This procedure would make the allocation of first choices a "pure matching" procedure, and the chance probability of four first choices coinciding with the correct correspondences would be one in 24. Thus a systematic violation of the instructions would increase the chance probability of the particular occurrence more than ten times.

Upon this dependence hypothesis there are a number of possible chance distributions conceivable. The most reasonable set of conditions, however, seems to be the following: It may be assumed that the judge limits his first choices in such a manner that he never designates one drawing as the first choice for more than one stimulus card. The "1" scores would then be distributed in a pure matching distribution. But if he utilizes this clear-cut system regarding first choices, his second, third, and fourth designations have no reason to be anything but randomly distributed. For if he is very sure of his first choices being correct, designation of further choices will not influence his score, so there is no point in systematizing them.

The hypothesis of dependent first choices would lead, however, to a chance score distribution somewhat different from the hypothesis of independence. The mean of this distribution is 10 as before, but the variance⁶ is 6. The standard deviation is therefore $\sqrt{6}$ or 2.449, slightly larger than that of the hypothesis of independence. CR's calculated upon this new standard deviation are smaller, but as may be seen in Table 3, the second matchings (the converse matching of responses to stimuli) by PH are still past the significance criterion. The total of all judgments gives a CR of 3.63, which is highly significant.

Hypothesis of Correlation

Another type of dependence must be investigated. Although the total CR's on the two previous hypotheses have been significant, it is evident that this total is based upon the deviation of four different sets of judgments of the same data. It is possible that in a second

⁶ The general formulas for the second moments of these distributions are, when n equals the number of drawings matched, $\frac{n(n^2 - 1)}{12}$ for the binomial, and $\frac{n(n^2 + 2)}{12}$ for the restricted dependent case. I am greatly indebted to Dr. J. A. Greenwood who developed the moments of the restricted distribution.

matching the subject will not make an entirely new judgment, but will be strongly inclined to designate the same first choices as in the first matching, at least to some extent; that is, there will be a positive correlation between his choices in the two matchings.

Furthermore, it is evident that a similarity that may determine the rating of a correspondence by one judge is equally available to a second judge. If the stimulus card pictured a circle, and the response drawing was a circle, then most reasonable judges would rate the correspondence as a first choice. The sum of a large number of judgments would give an extra-chance evaluation out of all proportion to the chance likelihood of two circles being associated.

It is necessary, then, to correct for this correlation. By using all the terms of the formula for the standard deviation of a sum:

$$\sigma_s^2 = \sigma_1^2 + \sigma_2^2 + 2r\sigma_1\sigma_2$$

the presumed dependence can be corrected for by substituting the observed correlations for r in the formula.

The observed correlation of the separate sets of judgments in the seven experiments are given in Table 5. They are predominantly

Table 5
INTERCORRELATIONS OF SUCCESSIVE SETS OF JUDGMENTS

| | Pa Pb* | Oa Ob | Pa Oa | Pb Ob | Pa Ob | Pb Oa |
|-------------|--------|-------|-------|-------|-------|-------|
| Exp. A..... | .75 | .70 | .60 | .60 | .80 | .60 |
| Exp. B..... | .70 | .75 | .60 | .60 | .50 | .55 |
| Exp. C..... | .30 | .60 | .40 | .20 | .20 | .20 |
| Exp. D..... | .35 | .40 | .05 | .05 | -.10 | .10 |
| Exp. E..... | .75 | .65 | .30 | .10 | .25 | .40 |
| Exp. F..... | .70 | .25 | .15 | .35 | .00 | .20 |
| Exp. G..... | .70 | .40 | .70 | .25 | .35 | .40 |

*The top legend indicates that the correlation is between PH's [P] first matching of stimuli to responses [a] to PH's [P] second matching of responses to stimuli [b]; the other participants' [O] first matching [a] to the other participants' [O] second matching [b]; and so on for the six possible inter-correlations.

positive, so their use in the formula increases the standard deviation and reduces the CR.

The total of PH's judgments upon the hypothesis of dependence and of correlation gives a CR, for the deviation, of 2.84, which is significant. The other subjects yield a CR of 1.23 which is not significant. The CR of the total is 2.42, which does not quite reach significance criterion.

The corrections for dependence and correlation provide a very severe test of the significance of the relatedness in the data, in that they constitute an over-strict null hypothesis. The resulting CR of 2.42 is close to, but does not exceed, a significant criterion of 2.5. But the *p*-value of the observed CR is .0072, which is lower than the reasonable probability criterion of .01 for significance.

The corrections used are over-strict because the hypothesis of dependence as it has been used here does not apply strictly to these data. Reference to Table 2 will show that the judges do not consistently distribute their first choices in such a way as to designate four distinct pairs of correspondences. (In the table such a distribution would be indicated by no two first choices occurring in the same row or the same column. In the first square of the table, two 1's appear in row a. In the square just below, two 1's appear in column B. In the fourth square of Experiment A, the first choices are systematically distributed to distinct pairs of correspondences.) The best statistical hypothesis to match what actually occurred would yield a variance somewhere between 5 and 6 in value; so an approximation to the true CR of these experiments would lie somewhere between the CR's observed upon the hypotheses of independence and dependence.

In the second place, control of correlation assumes that correlation leads to a spurious over-evaluation if uncontrolled. The assumption is only partially true. If there is an objective similarity between stimulus and response that leads to consistent rating by judges, this similarity may be chance, but it also may be evidential of an extra-sensory perception. A perfect reproduction of the stimulus cards by the responses would yield a high correlation of judgments, probably high enough that the scores of additional judges would not notably increase the CR evaluation. As a unit of evidence for ESP, the experiment would have a value not much greater than the CR of a single set of judgments. Yet the fact that the stimulus-response relation is so marked that it is consistently recognized by a number of judges is of considerable evidential weight. The assumption that all correlation is spurious excludes the common-sense weight of this evidence of consistency.

Thus a true CR would probably be considerably greater than the CR's computed upon the hypothesis of correlation and dependence.

However, such a correction is on the safe side, and should constitute a necessary part of the statistical evaluation of the experiments.

SOME IMPLICATIONS OF THE RESULTS

Inspection of the stimulus drawings and the percipients' responses shows that there is very little self-evident relation. However, when it is known that one belongs with another, some "relations" may be educed. For example, in Experiment A the response, apple, followed the contours of the stimulus, face (in fact, upon close measurement, to an astonishing degree of congruity), and the tree and hill carried out various part-form characteristics of the stimulus, leaf. In Experiment B the stimulus, wineglass, was associated with a cupboard, dish, and repeated references to things to eat. In Experiment C the form quality of the response to the anchor stimulus was strikingly similar, and so on. (See Plate 1.) But in all there are no examples of unequivocal reproduction customarily evaluated as successful "hits." The true content of the objective relation between stimulus and response is in no case clearly apparent.

The significant statistical results require the assumption of some extra-sensory mode of perception between the stimulus and response situations. Since the method is that usually characterized as general or undifferentiated ESP, wherein the stimulus may be either the card symbol or the mental process of an agent, or both, we can localize the ESP process no further than to say that it occurs between the stimulus and response situations. In the light of the results, it is interesting to note that the direction of the process is ambiguous as well. Inspection of Table 3 shows that PH's judgments were consistently superior to those of the other subjects. The difference is not statistically significant with so small a number of cases, but an inference of a true superiority should not be far wrong. The superiority in matching ability can scarcely be attributed to intellectual endowment, and certainly not to training. I have no intelligence record of PH, but upon report of his scholastic record would judge him to be approximately equal in potential intellectual level to the other subjects, who had all engaged in graduate university study. All of the other subjects had had some contact with the literature of ESP tests with drawings, and some had had considerable experience in the matching procedure itself. PH knew nothing of the free

drawing type of investigation. It would appear simply that PH in choosing from among the many possible reasons for matching two drawings had somewhat more effective guidance among his mental processes in hitting upon the correct cues.

But in these experiments PH was predominantly an agent; that is, he was part of the "stimulus situation." In only two of the seven experiments was he the percipient. Two implications are apparent. Either the content transmitted by ESP was derived more from his mental processes than from the stimulus object (that is, the ESP was predominantly telepathic); or PH as agent had some ESP impression of the response drawings which guided his later matching judgments. If the former is true, we should expect better results with ESP tests of this character than with clairvoyance tests in which no agent looked at the cards. If the latter is true, we can make no prediction of the results. But it would be possible to affirm that PH's performance was of the type we would expect from a "gifted subject."

Whether PH actually was a better agent than percipient is not possible to determine from the small amount of the material here. In Experiments A and B, he was paired with Russell as agent and percipient, and in Experiment C he was paired with me. If we take the total scores of all judgments of these experiments to characterize their relative "success," it is apparent that in the two in which PH was percipient (A and F) the total deviation from chance is just zero; whereas in the two in which PH was agent (B and C) the total deviation was -33. This apparent difference has a CR of only 2.17 on the hypothesis of dependence and correlation, which yields a suggestive but not significant probability of .02 that the difference is a true one.

A PERSONAL OBSERVATION

In Experiment C, I was the percipient. During it I had what might well have been an "experience" of ESP. Let us concede that the evidential value of Experiment C rests wholly upon its statistical evaluation (and alone it is not significantly different from chance), so that this report of "experience" is of speculative interest only.

My technique for promoting imagery was to close my eyes, watch for distinctive black forms in the periphery of the gray field, and to

interpret those forms into something to draw. At the outset of the second drawing the black form of two curved lines appeared distinctly, with other lines vaguely associated, and persisted to a remarkable degree. I tried in three attempts to "copy" them, giving the first an unsatisfactory interpretation—"umbrella." The persisting image finally disappeared, the two lines decreased in size and became antennae upon an insect which I characterized unsatisfactorily as an "ant or roach." (See Plate I.)

When the stimulus card with an anchor was presented to me, my experience was characterized in the remark, "That was what I saw!" The "recognition" was indistinguishable from any other experience of recognition of an object previously perceived. I emphasize this fact, because the obvious objection to the veridicality of such experience is that they may be motivationally determined illusions. If this was such an illusion, there was nothing to distinguish it from customary perceptual experience.

This personal experience may serve a general orientative purpose, however. In evaluating ESP research a great deal of attention has been paid to rigor of conditions, possibilities of error and sensory leakage, etc. At least in this instance the reason for matching the curved lines with the anchor stimulus card was wholly uninfluenced by any search for extraneous cues or attempt to systematize the scoring.

CONCLUSION

The method of preferential matching proposed here appears to be a useful addition to methods for the statistical evaluation of free drawing material in ESP tests.

Its sensitivity is attested by the fact that although in the data here considered there were no self-evident reproductions of stimuli in the responses, a small degree of relationship apparent to inspection is reflected in the measure of statistical significance.

Some degree of validity is suggested by the fact that a purported "good ESP subject" made statistically significant judgments as a participant in the tests. Since the investigation had the purpose of eliciting ESP ability from this subject, his significant results affirm the occurrence of ESP in this investigation.

The occurrence of ESP within the experiments is further attested by significant statistical differences of the results from those expected upon three reasonable chance hypotheses.

REFERENCES

1. CARINGTON, W. Experiments with paranormal cognition of drawings. *J. Parapsychol.*, 1940, 4, 1-117.
2. PRATT, J. G. The work of Dr. C. Hilton Rice in extra-sensory perception. *J. Parapsychol.*, 1937, 1, 239-259.
3. PRINCE, W. F. *The Sinclair Experiments Demonstrating Telepathy.* Boston: Boston Society for Psychic Research, 1932. Pp. 138.
4. SINCLAIR, U. *Mental Radio.* Pasadena: Upton Sinclair, 1930.

A CHECK ON SALIENCE RELATIONS IN ESP DATA

By J. G. PRATT, BETTY M. HUMPHREY, and J. B. RHINE

Parapsychology Laboratory, Duke University

ABSTRACT: It appears to make a difference what the position of a trial in the run is, as to the subject's success in identifying the cards in ESP tests, especially of the type called DT (calling down through the deck without removing the cards). This effect of position in the run leads to hit-patterning, in which certain parts of the run stand out above the rest as to rate of success in calling. For brevity, this standing-out is called "salience." Furthermore, the same effect has been found to hold for the subdivisions of the run, or segments, of five trials each; in fact, a relation has been found between the salience of the run as a whole and the salience of the segments in the same tests. This relation has been statistically significant in two researches that have already appeared in publication, and is being found in others to be reported later. It is important, therefore, to examine critically the way in which this result is obtained.

In order to test the method by which the relation between the salience of the segment and that of the run is measured, we have made an empirical check on the basis of the results of a research of 1,114 runs of DT tests reported earlier by Rhine as showing this relation to a significant degree. For this check the distribution of hits as they actually occurred in the twenty-five positions in the run were shuffled in a thoroughly random fashion, so that we had what should be expected to be a chance distribution. The salience for the segment and that for the run were computed in the customary manner, and the relation between the two measured as usual. The results were insignificant, and are ascribable to chance.

We conclude from this empirical check that it is not the method of analysis itself which has given the significant salience effects reported in the earlier studies.

INTRODUCTION

EXPERIMENTERS in ESP have lately been giving special attention to hit-distribution patterns in the run of twenty-five trials. In two papers that have appeared, one in September of last year (2) and one in this number (3), these hit-distribution patterns have been referred to under the general name of "saliency," or the degree to which the ends of the run stand out as compared to the middle, or vice versa. This saliency effect is of two kinds—that which applies to the run as a whole, and a similar effect discovered in parts of the run, or segments, consisting of five trials each.

The saliency of either run or segment is measured by a saliency ratio or SR, which is designated as SSR for the segment and RSR for the run. These SSR's and RSR's have been found to show a positive correlation as measured by the covariation statistic; in fact, the most striking evidence coming from the saliency studies lies in the significant covariation found between the SSR's and RSR's. Not only was this true of the first report on saliency mentioned, but in the second it is this feature that constitutes the sole basis for the conclusion that the results are not ascribable to chance. And further work now ready for publication is equally dependent upon the SSR-RSR covariation for its evidence of statistical significance. The evidence, then, is considerable that under certain experimental conditions the degree to which the SSR's and RSR's vary together is extra-chance.

But since the SSR and RSR are both computed from the same data, the question arises as to whether the two measures are mutually independent; that is, whether an increase or decrease in the one would be expected to influence the other. If there be interdependence between the two, it would be necessary to know its extent before any reliance could be placed upon the covariation CR's which measure the SSR-RSR interrelations.

From each Subseries or other grouping of results, there have been two hit-distributions made throughout the saliency studies. One is the distribution of the run in five subtotals, one for each segment of five trials. The discussion will be clearer if illustrated by the simple diagram of Figure 1, which represents a distribution of hits for a Subseries of runs of twenty-five trials broken into five segments of five trials each. In the diagram the distribution for the run is ob-

| | | SEGMENTS IN THE RUN | | | | | Total Hits | Dev. |
|------------|---|---------------------|-----|-----|-----|-----|------------|------|
| | | 1 | 2 | 3 | 4 | 5 | | |
| Trials | 1 | 26 | 27 | 31 | 29 | 29 | 142 | - 8 |
| | 2 | 29 | 27 | 26 | 35 | 30 | 147 | - 3 |
| | 3 | 29 | 33 | 30 | 29 | 29 | 150 | 0 |
| | 4 | 27 | 33 | 30 | 32 | 30 | 152 | + 2 |
| | 5 | 25 | 28 | 30 | 29 | 28 | 140 | -10 |
| Total Hits | | 136 | 148 | 147 | 154 | 146 | | |
| Dev. | | -14 | -2 | -3 | +4 | -4 | | |

FIG. 1. Diagram representing the distribution of hits for an actual Subseries of 30 runs of 25 trials per run. Each run is broken into 5 segments of 5 trials each. The 5 vertical columns represent the successive 5-trial segments of each of the 30 runs. The number in each cell represents the total number of hits in that position for the 30 runs. The RSR is computed from the deviations of the totals of the vertical columns; the SSR, from those of the horizontal columns.

tained by adding the hits in the vertical columns. The other distribution is that for the segment, and is obtained by adding across the horizontal columns, getting five distribution totals for the five corresponding positions in the segment. These two hit-distributions, each represented by five subtotals, are the figures from which the salience ratios are computed. A description of this computation is given in the first paper on terminal salience (2), and need not be repeated here in detail. Briefly, the distribution of deviations from chance is obtained from the hit-distribution; and for each of the five deviations, corresponding CR's are obtained which are squared to give five chi-squares, one for each of the five subtotals of the hit-distribution. Thus for the segment and for the run there are in each case five chi-squares for each subseries of results. The salience ratio or SR is the quotient of the sum of the first and fifth chi-squares divided by the sum of the second, third, and fourth. If there is no salience, terminal or middle, the SR is .667.

It may be seen now from this sketch that shifting the number of hits up and down the vertical columns will directly affect the SSR's,

which are measured horizontally, but will not affect the RSR's which are measured vertically. The converse would be equally true. Either enlarging or diminishing the SSR would require the shifting of deviation out toward the ends or in toward the middle of the vertical columns, according to whether the deviation is either positive or negative; but since the RSR is based, not upon any distribution within the column, but rather upon its total deviation, there could be no conceivable effect made by such internal shifting. Accordingly, then, it does not seem logically possible that there could be interdependence of the kind discussed; that is, where rise or fall of the SSR or RSR would automatically require rise or fall of the other.

However, in view of the increasingly great responsibility left upon the SSR-RSR covariation in studies already reported and about to appear, it seemed desirable that some empirical check be prepared, to give further support to the logic upon which the investigators have proceeded in their use of this method. One of us, JGP, suggested the analysis here reported and, with the collaboration of BMH, conducted the empirical test.¹

PROCEDURE

The plan consisted of taking the hit-distributions of the results, the 1,114 runs of DT tests already reported in the first saliency paper (2), and randomizing their order in a thorough-going fashion; that is, for each Subseries, which is the smallest grouping of data for which a distribution is available, the hit-totals for the twenty-five positions in the run were thoroughly redistributed in random order, so that they constituted presumably a chance distribution. This was done for the twenty-eight Subseries of the work reported in the first saliency paper.

From these scrambled hit-distributions, the proper groupings were made to obtain the hit-distribution for the segment in terms of five subtotals and the hit-distribution of the run in five corresponding subtotals. From these, the SSR's and RSR's were computed as usual, and the covariation between these presumably chance SR's was obtained in the established manner.

The same precautions were applied to this test as are customary

¹ The counsel of Dr. J. A. Greenwood, mathematical consultant of this Laboratory, is acknowledged with appreciation.

in the handling of ESP results themselves; that is, there were separate computations made by two different individuals, beginning with the transcription of the hit-distributions from the previously double-checked records, and including the randomization and the calculation of SR's and covariation CR's. The results represent the agreement of these independent analyses.

The method of randomization consisted of using a deck of twenty-five cards numbered consecutively from 1 to 25, shuffling it through a rotating shuffling machine, and cutting the shuffled deck by a number derived from throwing two dice. The order of the numbered cards as it then existed was taken as the order of redistribution for the Subseries concerned. A separate shuffling was done for each Subseries, making twenty-eight redistributions in all.

RESULTS

The SSR-RSR covariation was found to be insignificant for all four levels of SR's upon which the computation was made in the earlier salience work; that is, on the level of the smallest grouping, the Subseries, on the level of the Section SR's, the Series SR's, and the Division SR's. In order to understand these subdivisions the reader should consult Page 186 of the first salience report, and especially the diagram in Figure 1. It should be explained, however, that these four analyses are all made on the same data; the difference is in the size of the subdivision concerned, the Divisions being the largest groupings.

Table 1

COVARIATION BETWEEN SSR'S AND RSR'S FROM AN EXPERIMENTAL SERIES OF DT TESTS AND FROM AN EMPIRICAL CONTROL

| EXPERIMENTAL | | | | CONTROL | | | |
|-----------------|------|-------|-----|------------|-----|-------|-----|
| A. BY DIVISIONS | | | | | | | |
| Div. | SSR | d.f. | RSR | Div. | SSR | d.f. | RSR |
| Child | .80 | 24,36 | .84 | Child..... | .70 | 24,36 | .41 |
| Adult..... | 2.00 | 32,48 | .85 | Adult..... | .45 | 32,48 | .92 |

CR (cov.)=2.16 P=.02

CR (cov.)=-.95 P=insig.

B. BY SERIES

| Ser. | Div. | SSR | d.f. | RSR | Ser. | Div. | SSR | d.f. | RSR |
|---------------|------|------|-------|------|---------------|------|-----|-------|------|
| Interrupted.. | Ch. | 1.68 | 8,12 | 1.25 | Interrupted.. | Ch. | .82 | 8,12 | .66 |
| | Ad. | 3.73 | 12,18 | 1.57 | | Ad. | .22 | 12,18 | .72 |
| Uninterrupted | Ch. | .62 | 16,24 | .61 | Uninterrupted | Ch. | .66 | 16,24 | .32 |
| | Ad. | 1.40 | 20,30 | .62 | | Ad. | .64 | 20,30 | 1.09 |

CR (cov.)=3.37 P=.002

CR (cov.)=-.20 P=insig.

C. BY SECTIONS

| Sect. | Div. | SSR | d.f. | RSR | Sect. | Div. | SSR | d.f. | RSR |
|--------|------|-------|-------|------|--------|------|-----|-------|------|
| D..... | Ch. | .57 | 4,6 | 1.13 | D..... | Ch. | .92 | 4,6 | .72 |
| "..... | Ad. | 2.92 | 6,9 | 1.09 | "..... | Ad. | .23 | 6,9 | .72 |
| N..... | Ch. | 11.10 | 4,6 | 1.58 | N..... | Ch. | .53 | 4,6 | .42 |
| "..... | Ad. | 5.21 | 6,9 | 1.84 | "..... | Ad. | .20 | 6,9 | .72 |
| W..... | Ch. | .42 | 8,12 | .67 | W..... | Ch. | .54 | 8,12 | .57 |
| "..... | Ad. | 2.18 | 10,15 | 1.00 | "..... | Ad. | .67 | 10,15 | .80 |
| C..... | Ch. | .93 | 8,12 | .57 | C..... | Ch. | .75 | 8,12 | .18 |
| "..... | Ad. | .88 | 10,15 | .43 | "..... | Ad. | .64 | 10,15 | 1.52 |

CR (cov.)=3.21 P=.002

CR (cov.)=-.25 P=insig.

D. BY SUBSERIES*

| CHILD | | ADULT | | CHILD | | ADULT | |
|-------|------|-------|-------|-------|------|-------|-------|
| SSR | RSR | SSR | RSR | SSR | RSR | SSR | RSR |
| .40 | 1.43 | 12.36 | 1.78 | 1.63 | .46 | .26 | 2.75 |
| 10.73 | .93 | .13 | 1.43 | .42 | 1.64 | .37 | .20 |
| 41.15 | 1.16 | .26 | .05 | .42 | .50 | .03 | .02 |
| 5.00 | 6.28 | 16.94 | 9.25 | 1.45 | .09 | .15 | 1.86 |
| .18 | .17 | 51.88 | .90 | .02 | .89 | .12 | .52 |
| 3.34 | 1.66 | .47 | .29 | 1.96 | .46 | .28 | .06 |
| 1.31 | .50 | .54 | 1.68 | 6.09 | .12 | 1.23 | .72 |
| .06 | 4.47 | 5.00 | 1.49 | 1.01 | 1.20 | .12 | .10 |
| 1.59 | .11 | 2.56 | .02 | 5.92 | 1.25 | .22 | 1.59 |
| .18 | .06 | 3.74 | .05 | .61 | .28 | .69 | .92 |
| 2.88 | 1.54 | 1.96 | 16.45 | .51 | .06 | 3.70 | .75 |
| .97 | 2.40 | .59 | 1.22 | .30 | .09 | .65 | 14.58 |
| | | 13.00 | .07 | | | .14 | .29 |
| | | 5.62 | .45 | | | .17 | 1.68 |
| | | 1.87 | .03 | | | 1.07 | 1.41 |
| | | .16 | 1.42 | | | .46 | .44 |

CR (cov.)=1.19 P=insig.

CR (cov.)=.22 P=insig.

*The degrees of freedom for each Subseries are 2, 3.

The SSR's and RSR's of the experimental data themselves are presented in Table 1 on the left for comparison with the corresponding SR's obtained from the control data which are given on the right. It may be noted that not only are the control SR's on the whole much smaller than those of the experimental series, but they appear, even to inspection, much less positively related. By this we mean that SSR's and RSR's are more often on opposite sides of the no-salience point of .667. There is also much more terminal salience in the experimental series than in the control. It may be noted that whereas all the covariation CR's have thus far been positive, two of the control CR's are negative and two positive, while all are negligibly small, with insignificant probabilities in each case.

DISCUSSION

In view of the considerable number of separate subdivisions (representing 1,114 runs) randomized for the control series, there can hardly be further need for extending this control sample. Both logically and empirically it seems amply justified to dismiss from consideration any question of interdependence of such proportion as to warrant attention in these salience investigations.

SUMMARY

We have examined the question of whether the salience effect in the segment as measured by the SSR and the corresponding effect in the run as measured by the RSR are independent measures or are statistically interdependent. This examination was conducted because of the considerable importance being attached to SSR-RSR covariation in studies of salience.

Both on logical grounds and on the basis of an extensive empirical study, it is concluded that the SSR and RSR do not show evidence of interdependence on a chance theory. The empirical control series was produced by randomizing the hit-distributions of the results upon which an earlier salience study was based.

On the basis of this test, it must be concluded that the significance of the covariation between SSR's and RSR's in the experimental series cannot properly be attributed to a statistical artifact based on interdependence.

REFERENCES

1. GREENWOOD, J. A. The statistics of salience ratios. *J. Parapsychol.*, 1941, 5, 245-249.
2. RHINE, J. B. Terminal salience in ESP performance. *J. Parapsychol.*, 1941, 5, 183-244.
3. RHINE, J. B., and HUMPHREY, BETTY M. A transoceanic ESP experiment. *J. Parapsychol.*, [this number].

A TRANSOCEANIC ESP EXPERIMENT

By J. B. RHINE and BETTY M. HUMPHREY

*Parapsychology Laboratory, Department of Psychology
Duke University*

ABSTRACT: When, in 1939, word was received at this Laboratory that Dr. Karlo Marchesi, a physician of Zagreb, Jugoslavia, had produced significant scores in a lengthy series of ESP tests, it was proposed to him that he attempt to identify the cards in decks set up for the purpose in this Laboratory. He complied; and between August, 1939, and May, 1940, in three different series, attempted to identify the cards in a total of 353 ESP decks located in Durham. The score average was above the mean expected by chance, but not significantly so, except in the third series taken by itself.

However, we applied a method of evaluation recently introduced to this field of study, which is quite as reliable as the usual critical ratio method, but which is more complicated. This method utilizes the fact reported by Rhine last September that certain patterning of hit-distributions occurs in the run of twenty-five trials. It was found in this earlier work that these patterns for the run as a whole were reflected, as it were, in the segments, or five-trial portions within the run. As a matter of fact, it was the correlation between these patterns in the segment and the patterns in the run as a whole that constituted the major finding in the paper referred to. The method of determining whether or not the results were explainable by chance consisted in finding how much evidence there was of this relation between the hit-patterns of the segment and those of the run as a whole. This method is called "covariation between salience ratios," the term for the measure of the patterning. We found that the covariation of salience ratios for the segment and the run in the work with Dr. Marchesi gave a probability of .00014. This is a highly significant probability which clearly excludes the chance hypothesis.

It may not be clear at once how a series which averages but little

above chance can show significant hit patterns such as this work gives, but this is due to the fact that part of the pattern gives a negative deviation and part, positive. The patterning cuts across the "chance" line and the two trends of deviation cancel each other. Yet the patterns themselves (i.e., segment and run) are enough alike to exclude the chance hypothesis. The curves in Figure 1 give some idea of this relation, but, of course, the covariation results are the real basis for the conclusion.

Under the circumstances of the experiment, with 4,000 miles separating the cards and the subject who was trying to identify them, there can be no doubt that any knowledge shown would have to be extra-sensory. The only interpretation possible if the results are not explainable by chance or by sensory modes of perception is that an extra-sensory way of perceiving has been exercised in producing the results.

After Dr. Marchesi had acted as percipient in the first two of the three series, he proposed that a number of subjects in Durham attempt in like manner to identify the cards in decks set up in Zagreb. Ten subjects at Duke made a total of 1,000 runs, 10 each daily for 10 days, and these likewise gave a small positive but insignificant total deviation. When the scores of the Duke subjects were evaluated by the salience ratio and covariation method, the resulting probability was .12. This, too, was insignificant.

When the Marchesi data and the Duke data are evaluated together by means of the covariation of salience ratios of the run and the segment, the work as a whole gives a covariation CR of 3.11 and a probability of .0012, which is significant.

The fact that the hit-patterning or salience of the run is related to that of the segment is itself an interesting phenomenon, apart from the evidence it contributes to the occurrence of ESP itself. Such salience effects are being studied extensively, and a number of studies have been or will be submitted for publication, some of them directed to the study of the nature of this salience effect, and some utilizing it as a means of discovering ESP itself. Incidentally, there is also submitted with this report a note describing a series of tests which Dr. Marchesi conducted with the cards located nearby instead of at a great distance. These tests gave much higher averages than the

distance tests, though less striking salience relations. The fact, however, that they show positive salience relations, such as are shown by the distance tests, makes them of sufficient interest to be mentioned even though the conditions were less satisfactory. They are not presented, and in fact are not needed, as support of the case for ESP which the Marchesi distance experiment contributes.

This transatlantic experiment with Dr. Marchesi represents the longest distance which any systematic ESP investigation has, to our knowledge, ever utilized. This in itself is of some interest, especially since the most far-reaching conclusion of the nature of ESP concerns its place in the space-time universe.

INTRODUCTION¹

THE RESEARCH described in this report consists of transoceanic ESP tests conducted jointly by Dr. Karlo Marchesi of Zagreb, Jugoslavia, and one of the writers, JBR, in 1939 and 1940. The tests followed the card-identification method known as DT, in which the cards are left untouched throughout the test. The results are, by one method of analysis, ascribable to chance, but by another they are decidedly extra-chance in character. On the whole, the unusual conditions under which the experiments were conducted, and the analyses of the results themselves seem to warrant our bringing these data to the attention of students of this field.

The Main Percipient and the Setting of the Experiment

Dr. Marchesi had been in correspondence with JBR since January, 1939. In the course of this contact, he reported several experimental series of ESP tests in which significant results had been obtained. In some of these, he himself had acted as the percipient. When, in May of that year, JBR received from him the records of

¹The generous cooperation of Dr. Marchesi and of the ten American subjects who participated in these experiments is gratefully acknowledged. The assistance of Mr. Edmond P. Gibson in the recording of the cards and the checking of the records in the Laboratory is likewise very much appreciated, as is also the assistance of members of the Laboratory staff, and friends who have participated in the reading and correcting of this manuscript.

Due to the difficulties of correspondence in these times, Dr. Marchesi has not had an opportunity to see this manuscript before publication, and should not be held responsible for any of its conclusions.

200 runs of DT tests averaging 6.69 hits per run of 25 trials (a résumé of which is added below as a supplementary note to this report), he proposed that a transoceanic experiment be arranged, in which the cards would be set up in Durham and Dr. Marchesi would attempt to identify them from Zagreb. The response received was a cordial acceptance of the suggestion, and the first experiment was carried out in August of that year. Others followed in 1940, and later Dr. Marchesi suggested that a group of Duke percipients attempt to identify cards set up by him in Zagreb. A group of ten adults participated as subjects in this reverse arrangement.

Dr. Marchesi is a practicing physician, who, at the time of the experiments, held a government post in public health which he still retains. He is about forty-two years of age, and is of Italian, German, and Czech ancestry. He is the author of a book in the Croatian language entitled (as translated by him): *The Problem of Psychic Phenomena*. Dr. Marchesi states that although his medical training had inclined him to dismiss as incredible all such phenomena as that of extra-sensory perception, he was, as a result of his own experiences, and particularly his experiments, led to a distinctly favorable attitude toward the ESP hypothesis by the time the distance experiments with which we are concerned were begun.

Saliency

In an earlier report (10) JBR reviewed briefly the work which showed that when the DT test (calling down through the deck without removing the cards) was used, a higher rate of deviation from expectation was obtained at the beginning and end of the run than in the middle. The hit distribution, when plotted in five points for the run, accordingly gave a U-shaped curve. These curves were reported in 1934, and similar curves were later found by others. There were exceptions, but even with the exceptions there was the general feature that the ends of the runs tended to show greater deviation from expectation (in one direction or the other) than the middle section. This standing-out of the ends, or "terminal saliency" as it was called, appeared from the data available to be a general, though perhaps not invariable, effect of the DT test.

The DT work which was reported in a preceding article likewise gave a U-shaped curve, though in this case an inverted one.

This inversion, however, should be expected in view of the fact that the total deviation was negative. Pegram had obtained an inverted U-curve with the negative deviation obtained in her Low-Aim Series. With salient ends in a negative series, the inverted U-curve is a natural result.

The earlier experiment had been designed to vary the amount of terminal salience in the run, and also to induce a similar effect in the segments of the run that were produced by interruptions introduced at intervals of every five trials. As measured by the salience ratio (SR), a measure introduced for the evaluation of terminal salience, more of the latter was found in the interrupted runs than in the uninterrupted. This was true both of the runs as a whole and of the segments themselves. Thus, the standing-out of the ends of the run was increased by dividing the run into segments; and the segments themselves took on the character of small runs and showed even more terminal salience than the run as a whole.

Of primary importance was the significant relationship found between the SR's of the run and those of the segment, a fact which suggests that the salience shown in the one may be related to the other. Other significant relations statistically revealed between the SR's indicate that a general patterning of the distribution of hits is to be expected under the conditions of these DT tests, that this patterning is a matter of position of trials in the run and is in reality the reaction of the subject to the structure of the run as a whole. The segment-run relation between SR's was obtained by the covariation statistic, as described in an article by Greenwood (4) which accompanied the first salience report.

Salience as a Means of Exploring for ESP

In the preceding article on salience it was pointed out that significant evidence of salience was at the same time evidence of ESP; in fact, it was suggested that *the salience ratio and the interrelations between salience ratios are valid measures of the extra-chance character of ESP data—measures that might be used on experimental results in which more familiar methods would fail to show such extra-chance effects.* In this, and in other articles to follow, such indeed is the case.

The reasoning is as follows: If the hit-patterning which the SR (salience ratio) measures is not due to chance (that is, if there is an extra-chance distribution of successes throughout the run according to position or pattern) this distribution could presumably occur only if there were some cognitive connection between the percipient and the cards. Accordingly, to show that the results are extra-chance and hence are evidence of ESP, it need only be shown that there is significant salience or SR relations. That extra-chance patterning may occur in the absence of significant total scores is understandable when it is realized that this patterning may cut back and forth across the line of mean chance expectation, thus producing a cancellation effect in the total deviation of the series.

Significant salience effects have been found by three different methods: first, by the evaluation of the probability of the occurrence of an SR of given magnitude; second, by the covariation between corresponding SR's from two comparable subdivisions of like conditions in a research series—for example, in the preceding article this relation was computed for the Child and Adult Divisions; the third method consists in determining the covariation probabilities for the SSR-RSR relationship—the SSR's being the SR's of the segment of five trials, of which there are five in each run, and the RSR's representing the SR's of the run as a whole, in which the five segments correspond to the five trials in the segment.

Throughout the salience work reported thus far, it has been the covariation between the SSR and RSR that has contributed the best evidence for the extra-chance character of the results. This relation between the segment and the run has been a positive one throughout. It is the method applied in the present study.

The application of the SR-covariation technique to the Marchesi results came about in connection with a survey that is being made in this Laboratory of the occurrence of salience effects in DT and PDT (precognitive DT) tests for which the hit-distributions are on record or are obtainable from the records. The Marchesi records, which are based on DT tests, had not been published. (The experiments were, in fact, interrupted by the threat of war in Jugoslavia, but it is hoped they will be continued with the return of peace.) But because of their relevance to the salience inquiry it has seemed to us

advisable, in view of the international situation, to publish this interim report, dealing with that aspect of the material which is concerned with salience. This leaves for a later date, and perhaps for Dr. Marchesi himself, an eventual and more complete analysis and reporting of these, together with the other ESP experiments which he has conducted.

Distance and ESP

It may be seen from a review of the history of distance-ESP tests which JBR prepared for this JOURNAL in 1937 (7) that the Marchesi experiments involve a much longer distance than any others on record. There is, however, no reason, as far as the evidence goes, to expect any appreciable effect from longer distances than from shorter ones. While there are on record instances of a falling-off of success in ESP tests when the distance between percipient and the target cards (or agent, in telepathy tests) is extended, there are too many other instances of the lack of any such decline (with distance) to allow the conclusion that distance is a determining condition.

Added to the general evidence against any limiting influence of distance upon success as summarized in the 1937 article, there has appeared since that time a number of other striking cases that support a general conclusion. The cases of Warner (18) and of Riess (13, 14) involve only moderate distances, yet they are striking because of the high level of scoring attained by their subjects. The work of Carington (1) offers an interesting point because of the comparison of distance which is possible, since subjects in different parts of the world participated in his experiments. He mentioned that the group which happened to be located farthest away gave the best performance. This distance was from Durham, North Carolina, to England. The distance involved in the present report—approximately four thousand miles, from Durham to Zagreb—is therefore unique in extent.

PLAN OF THE REPORT

The experiments in which Dr. Marchesi acted as percipient will be called Section I, and those in which the Durham subjects participated will be called Section II. Section I will be further subdivided into four Subseries, consisting of 125, 90, 103, and 35 runs each,

given in chronological order. Section II consists of 10 Subseries, each of 100 runs, each Subseries representing the work of a single subject.

The subdivision of Section II calls for no explanation, but that of Section I should be described further. The runs of this Section were actually made in 3 sequences: the first, of 125, in August, 1939; the second, of 193 runs, in January and February, 1940; and the third, (prematurely interrupted) of 35 runs, in May, 1940. Since the Series of 193 runs was disproportionately large, it was broken into 2 Subseries, the first 90 having been conducted in January, and the remaining 103 in February. This separation was effected at the time of the checking-up of the experiments, which was approximately a year before the interest in salience arose in this laboratory. At the time the analyses were begun, we considered the question of keeping these January and February results separate, and decided in favor of doing so; first, on the grounds of comparable size of Subseries; and, second, on the grounds that even though the total deviations were slight, that of the January Subseries was positive and that of the February negative. The earlier terminal salience report had shown a tendency for a U-shaped curve of hit-distribution for positive series (2, 3, 6, 10), and an inverted U-curve for series with negative deviations (6). It is of some importance, as will be seen, that this division was made in advance of the analyses.

The Section results will be treated separately, and after the separate presentation, they will be combined.

PROCEDURES AND CONDITIONS

Since the conditions were essentially the same for all of the experiments, one description of the procedure will suffice. The day on which the tests would begin was agreed upon by correspondence in advance. Ten runs were to be done daily, and for that purpose ten target decks of plain ESP cards were shuffled, returned to their boxes, and laid out in a North-South line in the Parapsychology Laboratory at Duke, deck No. 1 being at the South. Dr. Marchesi was not informed of the details of their position, since he did not wish to give his attention to the details of localization.

The cards (already shuffled to break up original order) were

given at least five shuffles and a knife-cut before being set up for the first day. Thereafter daily for the test series the card order was recorded for each deck at the end of the exposure period, and the re-shuffling immediately followed the recording.

For Section I the cards were set up in Durham on the day preceding that on which the subject attempted to identify them. They were in place by 5:00 P. M. in Durham, which would be 11:00 P. M. in Zagreb. They remained there until 1:00 P. M. or later the following day, which gave Dr. Marchesi until 7:00 P. M. or later, Zagreb time, to record his trials. These were made, as a rule, soon after 8:00 A. M.

The standard commercial ESP Record Pads were used for recording. Each sheet has space for ten runs. Each run is broken into five segments of five trials each by the occurrence of double lines at intervals of five spaces in the column. At the end of the series, or when it was interrupted, Dr. Marchesi mailed copies of his records to this laboratory, where they were independently checked and re-checked. He was then informed of the results.

The same procedure was followed in the reversed arrangement of Section II, in which the subjects were in Durham, except that the ten decks of cards set up in Zagreb in Dr. Marchesi's office were arranged into rows of five each. In view of the considerable labor involved in the double-checking of one thousand runs which made up Section II, a copy of the card records were sent to Duke, and the checking was done by staff members here.

The same schedule of ten runs per day was followed in Section II as in Section I.

Methods of Analysis

The critical ratio method for the evaluation of total deviation of hits from expectation was applied as usual as each series was terminated, and the chi-square method of combination of CR's was used to obtain a general measure of the several series taken as a whole.

In the spring of 1941, about a year after the Marchesi work had been interrupted, the new method of evaluating the extra-chance character of certain types of ESP data, i.e., by the use of the salience ratio, was introduced. The manner in which the salience ratio is obtained and the several ways in which it may be evaluated have

been described in the earlier report, and it may be reasonably presumed now that those readers who would be interested in the present report have already become acquainted with the method. By way of review, however, a few words may be of assistance in recalling the essential outlines. The first step consists of making a hit-distribution for the twenty-five trials in the run for the group of data being evaluated. From this, two group distributions, each composed of five subtotals, are then obtained: the distribution of *the run* is made by grouping the hits of the first five trials as one total, those of the second as another, and so on, thus giving five subtotals. The hit-distribution *in the segment* is obtained by taking the total number of hits in the first, sixth, eleventh, sixteenth, and twenty-first trials, which gives the subtotal for the first trial in the segment. In a similar way the subtotals for the other four trials in the segment are found. Thus two distributions of the same hits are obtained: one in five subtotals for the segment of the run, and the other in five subtotals for the trials in the segment.

The deviation from expectation is then obtained from each of these sets of score-subtotals, a half-point dropped (to allow for discontinuity of the data; see p. 196, Vol. 5, this JOURNAL) and the critical ratio for each position is found. This is squared, giving a chi-square with one degree of freedom. From these sets of five chi-squares the salience ratios are computed by combining the first and fifth in each set and dividing their sum by the sum of the second, third, and fourth. The quotient is the salience ratio, or SR. The SR of the run is the RSR and that of the segment the SSR. The method of finding the probability of a given SR as well as the methods for measuring the relations between SR's (by covariation) are described in an article by Dr. Greenwood which accompanied the first terminal salience report (4). The chi-squares and the SR's computed from them are shown below in Tables 2 and 3.

RESULTS

Evaluated by Critical Ratio and Chi-Square Methods

The first Subseries of 125 runs gave a total number of hits below expectation by 11 points; the second, of 193 runs, came out exactly to expectation, although when divided into the January and February

Subseries of 90 and 103 runs respectively, the deviations became +16 and -16 respectively for the 2 Subseries. As it was decided in advance of the SR analysis that the series would be divided into 2 Subseries, it is so presented in Table 1, which contains the total results in terms of runs, deviation-CR's, and chi-squares.

Subseries 4 contained only 35 runs, since the threat of war in Yugoslavia interrupted Dr. Marchesi's activities as a subject. These

Table 1

TOTAL RESULTS EVALUATED BY DEVIATION-CR METHOD
Section I: Dr. Marchesi as Percipient

| Subseries | Period | Runs | Dev. | S. D. | C. R. | χ^2 |
|-----------|----------------|------|------|-------------|-------|----------|
| 1..... | August, 1939 | 125 | -11 | ± 22.36 | .49 | .24 |
| 2..... | January, 1940 | 90 | +16 | ± 18.97 | .84 | .71 |
| 3..... | February, 1940 | 103 | -16 | ± 20.29 | .79 | .62 |
| 4..... | May, 1940 | 35 | +34 | ± 11.83 | 2.87 | 8.24 |
| | Total..... | 353 | +23 | ± 37.57 | .61 | 9.81* |

$P (\chi^2=9.81; d. f.=4) = .044$

Section II: Duke Subjects

| Subseries | Subjects | Runs | Dev. | S. D. | C. R. | χ^2 |
|-----------|------------|-------|------|-------------|-------|----------|
| 1..... | LH | 100 | - 8 | ± 20.00 | .40 | .16 |
| 2..... | JGP | 100 | + 7 | " | .35 | .12 |
| 3..... | MB | 100 | +39 | " | 1.95 | 3.80 |
| 4..... | CES | 100 | +39 | " | 1.95 | 3.80 |
| 5..... | DHC | 100 | -19 | " | .95 | .90 |
| 6..... | LHG | 100 | + 5 | " | .25 | .06 |
| 7..... | BMS | 100 | 0 | " | .00 | .00 |
| 8..... | MS | 100 | - 1 | " | .05 | .00 |
| 9..... | FSC | 100 | +26 | " | 1.30 | 1.69 |
| 10..... | JLW | 100 | - 1 | " | .05 | .00 |
| | Total..... | 1,000 | +88 | ± 63.24 | 1.40 | 10.53* |

$P (\chi^2=10.53; d. f.=10) = .40$

| | | | | | |
|---------------------|-------|------|-------------|------|--------|
| Total Marchesi..... | 353 | +23 | ± 37.57 | .61 | 9.81* |
| Total Duke..... | 1,000 | +88 | ± 63.24 | 1.40 | 10.53* |
| Grand Total..... | 1,353 | +111 | ± 73.56 | 1.51 | 20.34* |

$P (\chi^2=20.34; d. f.=14) = .13$

*These chi-squares were obtained by addition of other chi-squares—not by squaring CR's, as were the others in this column.

35 runs gave a positive deviation of 34, with a CR of 2.87, for which the corresponding probability, allowing for both signs or directions of deviation, would be .004. This value needs, however, to be corrected for the fact that it is but one of 4 Subseries, and the P-value for Section I as a whole as evidenced by this Subseries would be .016.

Another, and perhaps a better, evaluation of this Section is obtained by means of chi-square combination of CR's. This, as may be seen from the table, gives a suggestive but insignificant probability of .04. It is necessary to conclude, then, that so far as the deviation CR method is concerned, the results are not significant.

It will be noted that the table presents also the results of Section II, and is divided into Subseries, each representing the work of a single subject. None of these is independently significant, nor is the total deviation itself. When the chi-square combination method is applied to all the 14 Subseries, it is likewise insignificant.

The Saliency-Ratio (Covariation) Method of Evaluation

Section 1. The principal method of utilizing SR's in the evaluation of the extra-chance character of a set of results consists in determining the covariation CR of the SSR-RSR relation; for example, the SSR of Subseries I is paired with the RSR of the Subseries. In like manner the SSR of Subseries II is paired with the RSR of the same Subseries. We have 4 such pairs of SSR's and RSR's for the covariation analysis of Section I. These 4 pairs are presented in Table 2 and by inspection it can be seen that they vary together to a remarkable degree. The results of the covariation treatment given

Table 2

SECTION I: MARCHESI RESULTS IN TERMS OF CHI-SQUARES AND SR'S

| Sub-series No. | Runs | Dev. | χ^2 IN THE SEGMENT | | | | | SSR | χ^2 IN THE RUN | | | | | RSR | |
|----------------|------|------|-------------------------|------|-------|------|------|------|---------------------|-------|------|------|------|-------|--|
| | | | 1 | 2 | 3 | 4 | 5 | | 1 | 2 | 3 | 4 | 5 | | |
| 1..... | 125 | -11 | .72 | 5.52 | .00 | 1.32 | .72 | .21 | .00 | .00 | 2.72 | .12 | .12 | | |
| 2..... | 90 | +16 | 2.53 | .03 | .08 | .17 | .08 | 9.32 | 9.06 | .42 | .00 | .28 | 1.25 | 14.73 | |
| 3..... | 103 | -16 | .37 | .03 | 2.92 | .15 | .15 | .17 | .00 | 7.29 | 2.56 | 1.61 | .25 | .02 | |
| 4..... | 35 | +34 | 1.51 | 3.92 | 7.51 | .22 | .08 | .14 | .22 | 9.73 | 1.51 | 3.24 | .22 | .03 | |
| Total... | 353 | +23 | 5.13 | 9.50 | 10.51 | 1.86 | 1.03 | .28 | 9.28 | 17.44 | 6.19 | 5.25 | 1.84 | .38 | |

in Table 4 show that this is a highly significant relationship, since it gives a covariation CR of 4.05, with a corresponding P of .00014. The fact that this covariation is positive also brings it into line with the significant positive covariation CR's in the previous salience report and in others now awaiting preparation.

Table 3

SECTION II: DUKE SUBJECTS' RESULTS IN TERMS OF CHI-SQUARES AND SR'S

| Sub-series | Subj. | Runs | Dev. | χ^2 IN THE SEGMENT | | | | | SSR | χ^2 IN THE RUN | | | | | RSR |
|------------|-------|-------|------|-------------------------|------|------|------|-------|-------|---------------------|------|------|------|-------|------|
| | | | | 1 | 2 | 3 | 4 | 5 | | 1 | 2 | 3 | 4 | 5 | |
| 1 | LH | 100 | - 8 | .71 | .53 | .03 | 1.66 | .00 | .32 | .38 | .25 | .53 | .08 | .08 | .53 |
| 2 | JGP | 100 | + 7 | .08 | .90 | .03 | 1.37 | .71 | .34 | 1.12 | .53 | .08 | 1.96 | .71 | .71 |
| 3 | LHG | 100 | + 5 | .08 | .03 | 1.12 | .38 | .38 | .30 | .00 | 1.12 | 1.66 | .15 | .53 | .18 |
| 4 | BMS | 100 | 0 | .15 | 4.28 | .25 | 3.84 | .00 | .02 | .03 | 1.37 | .53 | .08 | .90 | .47 |
| 5 | MB | 100 | +39 | 1.37 | .00 | .03 | .25 | 5.76 | 25.46 | 2.62 | .00 | .53 | 2.28 | 3.84 | 1.23 |
| 6 | CES | 100 | +39 | 1.12 | .25 | .90 | .00 | 2.28 | 2.96 | .03 | .71 | 1.37 | .90 | 1.96 | .67 |
| 7 | DHC | 100 | -19 | .90 | .38 | .15 | .00 | .90 | 3.40 | 1.66 | .53 | 1.37 | 1.37 | .71 | .72 |
| 8 | MS | 100 | - 1 | .38 | .00 | 1.12 | 1.96 | .08 | .15 | .25 | 1.5 | 1.12 | .03 | .00 | .19 |
| 9 | FBS | 100 | +27 | .90 | 1.37 | .03 | .03 | 1.37 | 1.59 | .71 | 3.42 | .00 | .03 | 4.28 | 1.45 |
| 10 | JLW | 100 | - 1 | .08 | .53 | .03 | .25 | .71 | .98 | .15 | .00 | 1.96 | 2.62 | .03 | .04 |
| Total | | 1,000 | +88 | 5.77 | 8.27 | 3.69 | 9.74 | 12.19 | .83 | 6.95 | 8.08 | 9.15 | 9.50 | 13.04 | .75 |

Table 4

COVARIATION BETWEEN SSR'S AND RSR'S

A. RESULTS AS A WHOLE

| Subdivision | No. of Pairs | Cov. CR | P |
|----------------|--------------|---------|--------|
| Sections..... | 2 | .86 | Insig. |
| Subseries..... | 14 | 3.03 | .0012 |

B. COMPARISON OF SECTIONS BY SUBSERIES

| Section | No. of Pairs | Cov. CR | P |
|------------|--------------|---------|--------|
| I..... | 4 | 4.05 | .00014 |
| II..... | 10 | 1.12 | .12 |
| Total..... | 14 | 3.03 | .0012 |

Because of the unusually close relation found between the SSR and RSR, more than ordinary interest attaches to the shape of the curve representing the distribution of the deviations themselves. The

deviation-distributions for the run as well as those of the segment are shown graphically in Figure 1, which represents the pooled deviations of the 4 Subseries of Section I; that is, all of the runs in distance tests in which Dr. Marchesi acted as percipient. There it will be seen that there is a remarkable paralleling of the segment and the run distributions. In spite of the fact that three of the four Subseries give middle salience, the total pooling produces curves showing some terminal salience, although the latter is not significant. Both curves

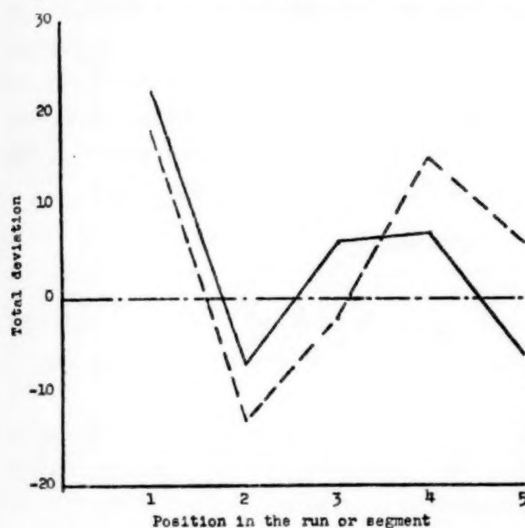


FIG. 1. Curves of total distribution of deviation of run (solid line) and of segment (broken line) in the distance DT tests with Dr. Marchesi as subject. 353 runs; deviation = + 23.

are U-shaped in the first 4 points, but the fifth point is dropped in both. This general shape is interesting in view of the long sequence of U-curves given by DT and PDT tests. (Compare also Figure 2 in which curves from other work by the same subject are offered. Especially notable is the similarity of the segmental curves [broken lines] in the two graphs.)

Section II. When the same method of evaluation is applied to the 10 Subseries of Section II, a positive but insignificant covariation is obtained. The covariation CR is 1.12, and the P-value is .12. For the details of the chi-square distribution, see Table 3.

Both Sections Combined. There is no statistical or logical requirement for the combination of the 2 series, although it is entirely legitimate to regard them as being, to some extent, parts of a larger

single exploration. Perhaps the most logical way of treating the 2 Sections in combination is to obtain a covariation CR for all 14 Sub-series representing both Sections as given in Table 4. This covariation CR is 3.03, and has a P-value of .0012. The corresponding CR for the Section level (+ .86) is of the same sign but is negligibly small.

Presumably, this figure of .0012 permits the conclusion that the work taken as a whole is significant, without attaching particular weight to the segregation of the effect of the work of Dr. Marchesi in Section I. It is of course statistically legitimate to treat his work as one of two Sections of the experiment. To do so would necessitate a correction for selection consisting of multiplying the P-value of .00014 by 2, making it .00028. Or again, a correction might be made on the basis of number of subjects; since there were 11 subjects, the P-value would be multiplied by 11, giving .0015.

DISCUSSION

Alternative Hypotheses

The following pertinent counter-hypotheses should be regarded before a conclusion is reached that ESP is evidenced in the experiment reported.

The usual safeguards of independent checking and computation have been employed, and the conditions of the testing obviate considerations of sensory perception's entering into the production of the results. The highly significant covariation CR obtained from the results with Dr. Marchesi as subject requires the rejection of the chance hypothesis.

The multiple-calling hypothesis is based upon the supposition that where more than one subject calls the same deck of cards, any similarity of symbol-patterns among the subjects might tend to affect the hit-patterning in some conceivable manner. It would be relevant here only to Section II and may therefore be omitted, since Section II is not involved in the conclusions.

The question of whether the SSR and RSR are, on a theory of chance, mutually independent, as the treatment in this paper supposes, has been discussed in the preceding paper on salience. There is no need to repeat that discussion here, except to say that it has been concluded that logically there is no ground for supposing inter-

dependence; rather, there is logical ground for supposing the two SR's to be independent. However, an empirical test has been made, and with the collaboration of Dr. J. G. Pratt, we are reporting it in this number of the JOURNAL. This test confirms the logical position taken in the earlier paper (II).

Altogether, there would appear to be no alternative to the conclusion that extra-sensory perception has occurred in these results, other than to find some weakness as yet unknown to us, either in the logic underlying the experiment, or else in the SR method of determining the extra-chance character of the results. Needless to say, both aspects of the work have had the best critical attention available to us in this laboratory.

Salience

The outstandingly high SSR-RSR covariation of Section I is the main feature of this study, but this ratio is by no means the only feature of interest. One novel aspect is that this significant covariation of SR's is obtained on the level of the Subseries ranging in size from 35 to 125 runs, whereas in the earlier work, covariation of SR's from Subseries of that size was not significant; that is, it required larger combinations than the Subseries represented to bring out any evidence of SR interrelation. Furthermore, the earlier report of salience was not based upon the work of a single subject, as this one is in Section I. The investigation of SR's of individual performers is not yet available for comparison with the Marchesi results, but it appears highly probable that his will be an exceptional if not unique individual study.

Another feature of the analysis of the results of Section I is the exceptional swing from extremely low SR's to extremely high ones, going from Subseries 1 to 2, and then through the equally phenomenal drop from Subseries 2 to 3.

It is plain that the Marchesi results showed *middle* salience rather than *terminal* in the August, February, and May Subseries. In January they showed strong terminal salience, mainly because the first position in both the run and the segment was high. But this terminal salience lasted only for the nine days in January, and gave way to striking middle salience that persisted likewise through the short series in May.

The chi-squares given in Table 2 show large fluctuations from Subseries to Subseries, especially in the three interior positions (2, 3, 4). For the run, these three positions together give a significant total chi-square of 28.88 that has, with 12 degrees of freedom, a P-value of .004. (It is interesting, too, that for the segment the corresponding value is .04.) These fluctuations of deviation, however, cancel each other in the pooling which produced Figure 1; the same must have occurred in both segment and run.

There is nothing in Dr. Marchesi's correspondence that gives a clue either to any change in conditions that might have caused the change in type of salience, or to the close similarity between the salience in the segment and that in the run. It is obvious that at least there is no simple, direct relation between SR's and the CR of the total deviation of the Subseries; for, while only one of the four Subseries (the fourth) has a significant positive deviation, it has about the same order of SR's as Subseries 1 and 3, which have small negative deviations.

One comment by the subject, however, may be of some relevance; namely, his express desire not to be given photographic detail as to the exact location of the cards. He wished to know only that they were "somewhere in space." His view was that a performance under those conditions would increase the value of the test if he succeeded in identifying the cards to a significant extent. It occurs to us that he might have been thrown back more upon his own imaginal structuring of the run, and less upon the form of the record sheet. This might have favored the concomitant structuring of both segment and run, and hence the SSR-RSR covariation. The more subjective his procedure, perhaps, the more prone he would be to such a great swing from middle to terminal salience and back, as occurred. But this is all very hypothetical as yet.

Distance and ESP Scores

In view of the earlier literature on ESP distance experiments, the question may arise as to whether these experiments may not indicate by their low score averages a retarding effect of distance, since, as was mentioned above, Dr. Marchesi had markedly better success in terms of high scores in DT tests conducted with the cards only a short distance away (see Note below). There are, however, many

other differences in the conditions of the two situations, and it cannot, therefore, be concluded that distance itself is responsible for the changed level of scoring. Particularly important may be the difference in familiarity with the local situation. While Dr. Marchesi courageously wished to attack the cards without specific knowledge of their location, it is quite conceivable that his unfamiliarity with their locus might have something to do with his rate of success. Furthermore, the great delay in finding out his success must have made the distance experiments much less lively and interesting than those conducted in his own office or home. The results of the distance tests afforded no day-to-day stimulation. Moreover, the state of the world through the period of the research was such that there was no certainty that the experiment would ever be finished, or if finished, that its results could ever be communicated to the subject. Other circumstances in the distance test could be pointed to on which hypotheses could be based to account amply for the low scores obtained, but there is no certainty to be had as to what actually caused the falling-off of deviation with transition to long-distance.

On the other hand, if we recognize, as we think the facts require, that there is unmistakable evidence of ESP, even though it is not in terms of total deviation, then it must be conceded that distance can hardly be supposed to have had any direct bearing upon the results, at least in the manner in which it is known to affect recognized physical energies; for, if the inverse square law were to be applied, a distance of 4,000 miles would surely reduce to absurdly minute intensity any energy that could have functioned in the close-up tests of Dr. Marchesi. We recognize that there is no clear and final way of determining whether distance might not be playing some part in such instances; no way, for that matter, of determining whether if a distance be sufficiently long (for example, expressible in terms of light-years) it might not diminish ESP capacity to the point of extinction. However, we may safely say that such energies as could at present be conceived to bear the transmission of the ESP symbols could not but be affected to the point of elimination by the distance of 4,000 miles. Also, the angle at which the cards lay (facing down), their close proximity to each other (25 to the quarter-inch in the box), must render them indistinguishable at a very much shorter

distance than from Durham to Zagreb. In short, considering all these difficulties and, in addition, the mountains and other barriers to any radiation of appropriate wave-length, we must reject all physical hypotheses as yet proposed. The fact, then, that any statistically significant effect was obtained stands out as perhaps only dramatizing a little more sharply the previously established conclusion (7, 12) that distance is not a limiting factor in ESP performance.

Time and Space Relations in ESP

It was the results of distance tests of ESP such as those reported here that raised the question whether time was a limiting condition to ESP performance. There have, by this time, been many researches contributing to the evidence that it is not; that is, that ESP may be precognitive (1, 5, 8, 9, 15, 16, 17). But before any systematic precognition experiments had been conducted here or elsewhere, it was asserted that the occurrence of precognition was logically inferable from the distance test results and related evidence (12; p. 291). This argument was elaborated by JBR in earlier publications (8, 9) and need not be reproduced here, except to say briefly that in a space-time world any process that is not affected by space limitations cannot be regulated by time, at least by the time which is dealt with by physics. Out of space, out of time, was the conclusion. But the argument is reversible, too, and we might say that the already considerable evidence for precognition supports in turn the hypothesis that ESP is not restricted by space. Thus the time and space experiments are, to this way of thinking, mutually confirmatory.

SUMMARY

Dr. Karlo Marchesi, a physician of Zagreb, Yugoslavia, acted as percipient in 353 runs of DT tests, with the decks of cards located in Durham, North Carolina. The results, as measured by the usual CR method, are only suggestive, but as measured by the SSR-RSR covariation statistic, a significant CR of 4.05, and a P-value of .00014 were obtained.

A similar series of tests of 1,000 runs was conducted with 10 subjects in Durham who attempted to identify cards in Zagreb. This series gave insignificant results by both measures; however, when

both sets of results were combined and evaluated by the covariation method, the result was significant ($P = .0012$).

The results are taken as evidence of extra-sensory perception, and, under the conditions, as evidence that distance is not an inhibiting factor in ESP performance.

NOTE ON DT TESTS CONDUCTED BY DR. MARCHESI ALONE

As we stated in the main body of the report, in May, 1939, JBR received the records of 200 runs of DT tests which Dr. Marchesi had conducted with himself as the subject. It was these results which led to the distance experiments. In sending these records to this laboratory, Dr. Marchesi mentioned only that he had made the tests by the DT method with a distance of 3 meters between himself and the cards, and that he was about to begin a series with 10 meters distance.

The 200 runs gave a total number of hits exceeding mean chance expectation by 279. This gives a critical ratio of 9.86, and an average score per run of 6.39.

Obviously, this series was distinctly not to be explained by the chance hypothesis, but due to the fact that the experiment was conducted in an informal manner—that is, with the subject himself playing the parts of both percipient and experimenter—there would be certain counter-hypotheses that could not be completely ruled out. There was no suggestion made that this series be published, and very probably Dr. Marchesi would concur in the view that it could not be taken as unquestionable evidence of ESP. He had evidently conducted the tests for his own satisfaction.

However, when the salience studies were made on the Marchesi distance DT records and this series of 200 runs was referred to again, it seemed in every respect proper material for a salience analysis. So far as could be determined, there was nothing in the conditions that could favor salience effects. Salience in the run and perhaps even, to a certain extent, in the segment might conceivably have been favored by the undue exposure of the top and bottom cards in the deck, but it could readily be determined if the first and twenty-fifth trials were solely responsible for any salience effect obtained.

Accordingly, the analysis was made. Both the segment and the run showed terminal salience, giving SR's respectively of 1.00 and

1.61. With only 2 and 3 degrees of freedom, these are not significantly covariant, but they are in agreement both "positive," and would of course be in line with the covariation found in the Marchesi distance tests.

The remarkable feature of the analysis, however, and the main reason for which the work is reviewed here, is in the deviation distributions for the run and the segment as represented in Figure 2. There it may be seen that the solid curve (for the run) is a very distinctly U-shaped curve, while that of the segment is a relatively close approximation. As mentioned above, the similarity between

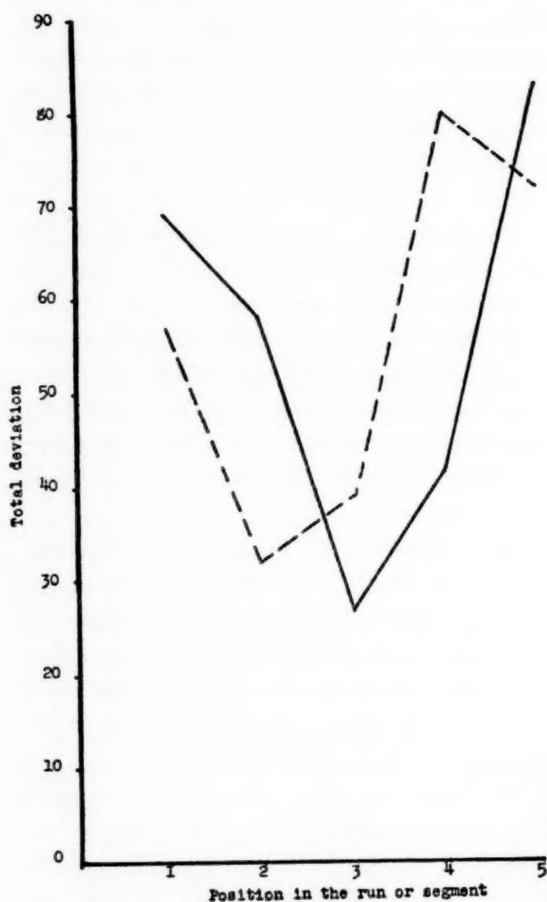


FIG. 2. Curves of distribution of deviation of run (solid line) and of segment (broken line) in Dr. Marchesi's DT tests of himself as subject. 200 runs; deviation = + 279.

this segmental curve and that given by Dr. Marchesi in his distance tests is noteworthy. In both, the first four points give a U, and the fifth point drops.

Although these are not statistical matters, they go far to reassure the critical reader on the question of the soundness of the experiment. There remains only the hypothesis that the most exposed cards, at the top and bottom of the deck, may have been identified by memory or visual perception. This question has been examined through the hit-frequency distribution in the run.

The following distribution was found for the 25 trials of the run: 57, 56, 42, 62, 52; 46, 50, 46, 60, 56; 49, 33, 51, 50, 44; 54, 41, 51, 51, 45; 51, 51, 49, 57, 75. From this it can be seen that the first trial is not particularly outstanding. The second trial is only one point below it and the fourth is five above it. It is only three points above the average for the first segment. In fact, the first trial could be reduced to the average for the first five without affecting noticeably the shape of the curve in Figure 2. The twenty-fifth trial has the highest frequency in the run, but, even so, the last place in the segment (see Figure 2) is not the highest. Furthermore, the run-distribution would still give a distinctly U-shaped curve, even if the twenty-fifth trial were reduced to the average of the other four trials in the last segment. These observations suffice to show that the shape of the curves in Figure 2 is not dependent appreciably upon the first and twenty-fifth trials, and accordingly it cannot be supposed that accidental observation of the top and bottom cards of the deck determined the nature of the distribution.

Therefore, it seems justifiable to add this series to the growing list of DT investigations (as well as PDT investigations still in manuscript form) that have yielded U-curves as well as interesting, if not significant, SR's and SR relations.

REFERENCES

1. CARINGTON, W. Experiments on the paranormal cognition of drawings. *Proc. S. P. R.*, 1940, **46**, 34-151; also *J. Parapsychol.*, 1940, **4**, 1-129.
2. CLARK, C. C., and SHARP, VERNON. Group tests for extra-sensory perception. *J. Parapsychol.*, 1937, **1**, 123-142.
3. GIBSON, EDMOND P. A study of comparative performance in several ESP procedures. *J. Parapsychol.*, 1937, **1**, 264-275.

4. GREENWOOD, J. A. The statistics of salience ratios. *J. Parapsychol.*, 1941, 5, 245-249.
5. HUTCHINSON, L. Variations of time intervals in pre-shuffle card-calling tests. *J. Parapsychol.*, 1940, 4, 249-270.
6. PEGRAM, MARGARET H. Some psychological relations of extra-sensory perception. *J. Parapsychol.*, 1937, 1, 191-205.
7. RHINE, J. B. The effect of distance in ESP tests. *J. Parapsychol.*, 1937, 1, 172-184.
8. ———. Experiments bearing on the precognition hypothesis. *J. Parapsychol.*, 1938, 2, 38-54.
9. ———. Experiments bearing upon the precognition hypothesis. III. Mechanically selected cards. *J. Parapsychol.*, 1941, 5, 1-57.
10. ———. Extra-sensory Perception. Boston: Bruce Humphries, 1934. Pp. 169.
11. ———. Terminal salience in ESP performance. *J. Parapsychol.*, 1941, 5, 183-242.
12. RHINE, J. B., PRATT, J. G., STUART, C. E., and SMITH, B. M., with GREENWOOD, J. A. Extra-Sensory Perception After Sixty Years. New York: Holt, 1940. Pp. 463.
13. RIESS, B. F. A case of high scores in card-guessing at a distance. *J. Parapsychol.*, 1937, 1, 260-263.
14. ———. Further data from a case of high scores in card-guessing. *J. Parapsychol.*, 1939, 3, 79-84.
15. SALTMARSH, H. F. Report on cases of apparent precognition. *Proc. S. P. R.*, 1934, 42, 49-103.
16. STUART, C. E. An analysis to determine a test predictive of extra-chance scoring in card-calling tests. *J. Parapsychol.*, 1941, 5, 99-137.
17. TYRRELL, G. N. M. Further research in extra-sensory perception. *Proc. S. P. R.*, 1936, 44, 99-168.
18. WARNER, LUCIEN. A test case. *J. Parapsychol.*, 1937, 1, 234-238.

GLOSSARY

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

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

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

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

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

CALL v.: To attempt to identify a *target* or *stimulus object* (or mental state of an *agent* in *telepathy*).

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

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

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

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

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

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*. (Occasionally the square of a simple *CR* may be used as chi-square.)

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

COVARIATION: Correlation evaluated in terms of theoretical means and *standard deviations*.

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

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

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

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

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

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

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

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

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

EXTRA-CHANCE: Not due to *chance* alone.

FREE MATERIAL: *Stimulus objects* that are not limited to a known number of categories.

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

HIT: The correct correspondence of a *subject's call* or response with a *stimulus card* or *object*.

HIT FREQUENCY DISTRIBUTION: The grouping of the total *hits* in a *series* of *runs* with respect to their original position in the *run*.

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

MATCHING: A form of *calling* in which a *target card* is placed opposite the *key card* which the *subject* selects to identify it. Also, in the evaluation of *free material*, the act of a judge in identifying a given *response* with a *stimulus object*.

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

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

PERCIPIENT: The person who makes the *calls* in a test situation.

- P (PROBABILITY):** A mathematical estimate of the expected relative frequency of a given event if chance alone were operative.
- RESPONSE:** The act of the *subject* in attempting to identify the *stimulus object*.
- RUN:** A succession of *trials*, usually the *calling* of a *deck* of 25 *ESP cards* or symbols.
- RUN SALIENCE RATIO (RSR):** A measure of *salience* within the *run*.
- SALIENCE:** The relation of rate of success in the end *segments* of the *run* to that of the middle *segments*; also the relation of the rate of success in the end *trials* of the *segment* to that of the middle *trials*.
- TERMINAL SALIENCE:** A higher rate of *deviation* in the end *segments* of the *run* (or in the end *trials* of the *segment*) than in the middle *segments* (or *trials*).
- MIDDLE SALIENCE:** A higher rate of *deviation* in the middle *segments* of the *run* (or in the middle *trials* of the *segment*) than in the end *segments* (or *trials*).
- SALIENCE RATIO (SR):** A measure of the relation of the rate of success in the end *segments* of the *run* (or in the end *trials* of the *segment*) and that of the middle *segments* (or *trials*). (For details of the manner of obtaining SR's, see Vol. 5, pp. 193-195.)
- SCORE:** The number of hits made in one *run*.
- TOTAL SCORE:** *Score* of any number of *runs*.
- AVERAGE SCORE:** *Total score* divided by number of *runs*.
- SCREEN:** An opaque barrier used between the *subject* and the card or *agent*. The main types of screens are illustrated in this journal on their first introduction in print.
- SEGMENT:** One of the five consecutive sets of five *calls* in a *run* of 25 *trials*. The first five *calls* would constitute the first *segment*; the second five, the second, etc.
- SEGMENTAL SALIENCE RATIO (SSR):** A measure of *salience* within the *segments* of the *run*.
- SERIES:** Several *runs* that are grouped in accordance with a stated principle.
- 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 2.5 or greater.
- STANDARD DEVIATION:** The theoretical root mean square of the *deviations*. For *ESP cards*, $SD = 2\sqrt{\text{no. of runs}}$.
- STIMULUS OBJECT:** The *ESP card* or drawing or other object, some identifying characteristic of which is to be apprehended by the *subject*.

STM (SCREENED TOUCH MATCHING): The technique in which the *subject* makes his *call* by pointing to one of five positions or exposed *symbols* under a special *screen*. The experimenter places the *target card* so designated in the position pointed to. The *screen* blocks all vision by the *subject* of the *cards* and their manipulation by the experimenter.

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

TARGET: *Stimulus object*.

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.

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

TRIAL: A single attempt to identify a *stimulus object*.