The

Journal *of* Parapsychology

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

Volume 6 SEPTEMBER, 1942 Number 3

Contents

PAGE

Editorial: Hypnotism, "Graduate" of Parapsychology . . . 159

Exceptional Scores in ESP Tests and the Conditions
I. The Case of Lillian 164
MARGARET PEGRAM REEVES and J. B. RHINE

The ESP Controversy 174
DOROTHY H. POPE and J. G. PRATT

A Confirmatory Study of Salience in Precognition Tests . . . 190
BETTY M. HUMPHREY and J. B. RHINE

A Second Study of the Effect of Tempo Rates of Matching . 220
CHARLES E. STUART and BURKE M. SMITH

Glossary 232

DUKE UNIVERSITY PRESS
DURHAM, N. C.

The Journal *of* Parapsychology

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 SEPTEMBER, 1942 Number 3

EDITORIAL:

HYPNOTISM, "GRADUATE" OF PARAPSYCHOLOGY

THERE HAVE BEEN a few statements expressed recently in print, implying some question of the status of extra-sensory perception because it has been criticized by "strict scientists" and "qualified psychologists." This situation provokes a bit of reminiscence about a field of phenomena that once belonged to parapsychology (or psychical research) and has since become accepted psychology—the field that first went under the name of "mesmerism" and is now known as "hypnotism."

Perhaps few persons today realize that hypnotism, generally accepted as it now is by psychologists, was originally one of the cardinal topics of investigation by the Society for Psychical Research back in the days of its founding in the '80's, long before hypnotism was a legitimate subject of study in academic psychology. The statement of the "Objects of the Society" expressly indicates this policy: "The following subjects have been entrusted to special Committees: 1. . . . . 2. The study of hypnotism, and the forms of so-called mesmeric trance, with its alleged insensibility to pain; clairvoyance and other allied phenomena." Moreover, many, if not most, of the better-known names in the history of hypnosis during the 80's and 90's and the early quarter of this century appear on the rolls of this society or in its publications. Among them are Richet, Janet, Schrenck-Notzing, Bramwell, Lloyd Tucker, Edmund Gurney, Frederic Myers, and a little later, T. W. Mitchell and McDougall.

The close association between hynosis and phenomena which

are still regarded as *para*psychological had been recognized for a long time before the date mentioned. Mesmer and his earlier followers encountered numerous instances in their mesmeric performances of what later came to be called "telepathy" and "clairvoyance." In the demonstration of hypnosis at a distance by French physicians, among them Janet and Gibert, during the last quarter of the 19th century, telepathy apparently was involved, and "traveling clairvoyance" experiments were a part of the hypnotic repertoire. Barrett's now famous report before the British Association for the Advancement of Science at Edinburgh in 1876 was concerned with both hypnotism and telepathy. Richet's experiments on clairvoyance with Léonie, the Sidgwick and Johnson research in telepathy, Backman's work in Sweden, and still others appeared to combine hypnosis and some extra-sensory mode of knowledge.

In fact, the twin phenomena of hypnosis and extra-sensory perception appeared on the scientific stage together during the last century. Neither one was recognized or studied by the academic psychology of that day. The two effects were believed to be produced only in rare individuals, neither was understood, and both were difficult of controlled study and verification. Hence, any systematic research concerning them was an irregular, unorthodox endeavor—not to be pursued by a university but by a more informal research society.

But in spite of this informal beginning, the phenomena of hypnosis have, during the last quarter century, become relatively well established in psychological circles and have done so much more rapidly than extra-sensory perception. With the coming to Harvard of Professor McDougall in 1920, hypnotism was introduced, not only to Harvard but to American psychology—indeed, we may add, to the academic world as well. Research both in hypnosis and in extra-sensory perception was done at Harvard under Professor McDougall. The former subject took hold firmly and is today "accepted," while the latter is only slowly and painfully gaining ground, with great labor and after much battling over fine points.

What is the reason for this great difference? One obvious factor is the easy repeatability on demand which goes with hypnosis. Once a person is hypnotized, he can be re-hypnotized over and over. No subject has ever been so regularly responsive in tests for extra-

sensory perception—hence the need for a statistical evaluation for the discovery of the latter. On the other hand, no statistics has ever been needed for the demonstration of hypnosis. This fact is enough in itself to account for the relatively more rapid acceptance of hypnotism as compared to extra-sensory perception. Reasoning from statistical results may be sound enough, but it is more intricate. There is, indeed, more danger of fallacy in such reasoning, but many of those who are not statisticians (which naturally includes a large number of psychologists) have an exaggerated conception of this danger and tend to distrust ESP statistics simply because the results support ESP. If, then, the phenomenon is one that can be established only by statistical methods (and there are many such in nature), we simply have to give these people more time. We cannot change the phenomenon.

It is probably not because hypnosis is "psychologically understandable" that it is more widely recognized. While it may, and we think does, appear to lend itself more readily to hypothetical explanation than does extra-sensory perception, there has been no thoroughgoing explanation of hypnosis that has had any legitimate scientific status to date. Animal magnetism, odic force, hyper-suggestibility, artificial sleep, and many other hypotheses have been offered, but they are either wholly untested notions in themselves or relatively empty names. Hypnosis, like ESP, is largely a mystery. For example, the hypnotized subject's freedom from pain when his unanaesthetized nerves are being cut is, like the identification of an unsensed object, something still to be accounted for.

We recognize that the "stop-gap" explanation of hypnosis as being hyper-suggestibility may have "naturalized" it, made it more plausible, for many students of psychology. But what, indeed, is hyper-suggestibility? The phenomena of hypnosis had been occurring for a hundred years before they became of psychological interest (except to medical psychology) on the basis of the "suggestion" hypothesis. And it may similarly be necessary for the rank and file of academic psychologists to have a plausible *hypothesis* of ESP, however speculative, to help bridge the gap of understanding before they can accept this finding. Of course, an untested hypothesis is only a half-way point in science—sometimes a poor one to stop at—

but it may have great bearing on the social side of research, which none of us can ignore.

One point seems to stand out from this history of the transition of hypnotism from psychical research to psychology proper; it is the fact that the disapproval of one's fellow-scientists, however they may be identified as "competent," "qualified," "leading," or what-not, may completely fail to indicate the ultimate value of a pioneer finding. This is particularly true of the school-ridden profession of psychology. Lay writers, especially, often appear not to realize that the mere fact of disagreement over a question in psychology tells practically nothing of the final truth of the issue at stake—that dissension is the rule, as yet, in that branch of inquiry. For example, Strömberg is apparently hesitant to accept the experimental evidence of telepathy because "several psychologists" have criticized "Rhine's conclusions"; and Rufus Jones states that "telepathy at best is a rare occurrence, still questioned by strict scientists, and at the present stage of verification a dubious ground of support." These writers are making wrong assumptions about the importance of agreement among psychologists. It would help them very much in seeing the position of ESP today in its proper light if they could realize that less than a generation ago almost all academic psychologists would doubtless have said that hypnosis is a fake act sometimes put on by vaudeville performers and that it has no place in scientific psychology. The same minds that were long closed to a consideration of the evidence for hypnotic phenomena would today be the first to refuse to look at the evidence for ESP.

However, if parapsychology continues to succeed, there may be others of its problems which will "graduate" into recognized lawful principles that cannot be rejected by general psychology, as hypnotism has but recently done. The parapsychology of today promises well to contribute its share to the psychology of tomorrow. But no one who knows the profession should expect unanimity on this point from our fellow-psychologists of the present.

* * * * *

Controversies in science, like wars, are wasteful procedures and seldom the best way to settle issues. But they are not entirely fruitless, and the battle over ESP is such an instance. The review of the

"ESP controversy" in this number is an attempt to show what the criticism has done to the research—how it has helped and how it has hurt it. Suggested reading for future critics, as well as future experimenters!

Only those few observers who have stood in the presence of the phenomenon of a perfect score of twenty-five in an ESP test can appreciate *adequately* the exhiliration experienced by Miss Pegram (now Mrs. Reeves) on the occasion described in "The Case of Lillian." Can explorers in this field catch, in or between the lines, the knack or insight that enabled Lillian to reach the heights of ESP proficiency?

During the last ten years, precognition—an hypothesis that appears to be all at odds with the present-day scientific world view—has become a topic of research in a number of places: in Holland, in England, and in America. While a number of other revolutions have been under way in the social and political order, persistent experimental efforts have been continued toward the crucial testing of the precognition hypothesis. The total cumulative effect of the findings is becoming weighty. In this issue, there follows a confirmatory sequel to the report of Rhine in the last number. Humphrey and Rhine now submit evidence that longer delay (ten days) in checking does not reduce success in predictions even when the subjects do not know which test will have an immediate and which a delayed check.

And the end of such confirmatory reports is not yet!

EXCEPTIONAL SCORES IN ESP TESTS AND THE CONDITIONS

I. The Case of Lillian

By MARGARET PEGRAM REEVES and J. B. RHINE

Undoubtedly there are other psychological phenomena which are quite as little understood as extra-sensory perception—for example, the genius of Beethoven or Shakespeare—but few, perhaps, on which so little fundamental understanding has been achieved in proportion to the amount of inquiry made. While this lack of understanding is testimony to the difficulty and obscurity of the problem of the nature of ESP, it should not be confused, as it often is, with the question of the occurrence of that process. Briefly, it is not correct to reject simply because we cannot explain.

On the other hand, we recognize clearly that it greatly facilitates acceptance for most people to have a ready-made explanation to garnish strange and, at first sight, unpalatable facts. Such a realization spurs the ESP investigator to a number of lines of activity which are not his primary experimental concern. First, to the encouragement of speculative and hypothetical discussions of the nature of ESP. Second, to the exploration of all possible similarities and analogies between ESP and other cognitive processes. Third—and here we come to the orientation of this particular article—to the consideration of every available report of research that may yield clues to the nature of ESP, whether or not there has been complete experimental perfection in conducting it.

To put the matter more bluntly, we are going to assume that the great mass of accumulated evidence from widely confirmed experiments has established the case for the occurrence of ESP and that accordingly its *possibility*, at least, in any given instance may be taken for granted. In the case of any experiment, then, in which there was some precaution omitted which has later come to be re-

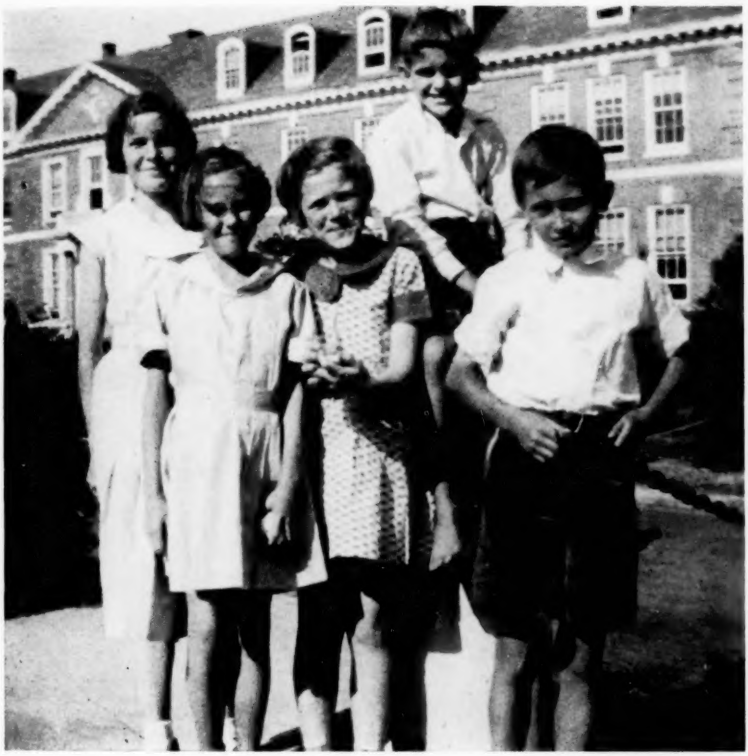

FIG. 1.—Group of Children from Wright Refuge.
Lillian is the central figure.

garded as essential to a fully adequate ESP research, we shall feel free to reconsider the case in this new light, the objective being to discover not primarily whether ESP occurred, but if it is reasonably likely that it did, *what conditions contributed* to its occurrence.

Especially do we wish, in attacking this general question of how ESP occurs and under what circumstances, to obtain the effect itself in its most striking manifestations. Many of these, it is true, have occurred under excellent conditions of safeguarding, such as those of the Pearce-Pratt series (3), the instance of Riess's subject (4), Warner's test case (5), and the distance telepathy series at Duke known as the Junaluska series (2: p. 104); but it is obvious that students and interpreters of ESP results do not have nearly enough factual material from which to draw insightful conclusions as to the nature and control of the effect under investigation. With this situation in mind, we propose to recover from the material which was withheld from publication because of some minor experimental deficiency, a number of case studies such as the one offered in the following pages. The guiding principles will be: (1) Were the experimental conditions good enough for a *general* psychological research? (2) Does the case offer something of apparent value bearing on conditions favoring or hindering ESP performance? We think the case of Lillian meets these requirements.

The Subject

Lillian was a normally intelligent child of nine years at the time of the ESP tests described. She is the central figure in the photograph reproduced on preceding page. At the time the tests were given her, she was staying at a children's home, the Wright Refuge, in Durham, North Carolina. Her stay there was temporary, pending adjustment of the home situation. She was one of a large family of children. Some conception of her maturity, her attitude toward the tests, and her relation with the experimenter can be had by reading the uncorrected letter reproduced below on page 170.

The Experimental Background

In the spring of 1936, one of us (MPR) was engaged in a series of experiments in which a number of types of ESP tests were being compared—five of them in all. These tests were given to a group

of 21 children at the Wright Refuge, a group composed of 12 girls and 9 boys, ranging in age from 6 to 13 years. The five test procedures compared were those known as OM, BM, DT, GESP, and ESP shuffle. The results have not been published hitherto, except for a brief review appearing in *Extra-Sensory Perception After Sixty Years* (p. 277).

The results of this work taken in its entirety are statistically significant, and show a number of interesting relations in addition to that which is selected for presentation and discussion here. Due, however, to the fact that extreme emphasis upon the need for independent recording was introduced about that time, and the fact that MPR's work did not involve such recording since she was the sole observer, it was deemed advisable to withhold this work from publication until the case for ESP was strong enough to warrant consideration of work done under more or less normal precautions of experimental psychology.

It is our purpose here to review the conditions of this research only inasfar as they are essential to the special performance of Lillian on the one ESP procedure which is concerned. This procedure is the OM or open-matching procedure, and it should be frankly recognized as not secure against the possibility of error. It consists of the following steps: The experimenter places five key-cards face upward in a row on a table. These represent the five suits of the ESP deck. The subject to be tested is given a deck of cards which has been shuffled and cut by the experimenter. The deck was made up in this Laboratory from specially selected die-cut stock with the symbols hand-stamped on them to prevent marking or warping. The cards were carefully inspected and the decks were re-examined for marks or cues before they were put to use.

The subject was instructed to think which of the key-cards the top card of the deck would match. She was then to place this top card opposite the identifying key-card, still keeping it face down. This procedure was to be continued until all the 25 cards of the deck were distributed opposite the key-cards. Then the experimenter turned the cards over, and she, together with the subject and other children present, counted the number of hits. No record was made of the card distributions—only the total score was entered in the experimenter's record book.

For the most part, the experimenting was done outdoors on the playground, the experimenter sitting on the ground before a very low table with the subjects sitting just opposite. Most of the time there were other children nearby. Some attempt was made to keep them at a specified distance from the subjects to avoid distraction, but this was not always successful. For the most part, the presence of the other children did not appear to distract the subjects. The children were accustomed to being together at all times.

The children were given candy rewards throughout the experiment. As each child finished his experimental session for the day, he was given a piece of candy for "trying." A different and more desirable piece of candy was given in addition as a prize for a score of 7 or more hits per run. A score of 12 received two prize pieces. The children were told more or less playfully that if anyone got a score of 25, he would receive 50 cents.

Lillian's Exceptional Performance

In general, Lillian's performance under the various test conditions placed her in the group of the four best subjects. On two methods, GESP and OM, she was the highest. Over the entire 100 runs which she made for all five conditions, she averaged 5.91 hits per 25, which is the highest of the group, the next being 5.70.

Lillian's performance was in general somewhat higher as the experiment continued as may be indicated by the averages of the five equal sections into which the experiment may be divided: 5.34, 5.36, 5.26, 6.10, and 6.05. Against this general background there stand out the striking instances of two runs which we propose to present and discuss in some detail. First, we will quote from the report of MPR submitted shortly after the performance took place, and based upon the notes made at the time:

"Lillian had two very remarkable scores on OM, which, naturally, made her scores on this condition disproportionately high. During one experimental session with her she was given a deck of cards which she had not used before (although the cards had been used by other subjects); she gave the cards a cut, placed them, and scored only three hits. Then she called the cards DT with five hits; the deck was shuffled by the experimenter and given to her again for OM. Again she cut the cards once, and placed them, this

time scoring twenty-three hits. She was pleased, but not especially elated over this performance. Rather, she regretted the misplacement of the other two cards.

"The next experimental session she was given a different deck of cards. First she scored three on OM, eight on BM, and twenty-five on OM. The experimenter had shuffled the cards between each run, and was naturally observing the performance exceptionally carefully after the first high run of the previous period. There were also two of the older children observing her, and although their observation is perhaps of less value, they reported they saw nothing suspicious. Lillian did not appear to be looking at the backs of the cards so that she could have seen the symbols, and I was the only person in front of her who could have seen the cards had they been held at a proper angle for this, which they were not. Further, the table on which she was working was dull and unpolished, so that there could have been no possibility of seeing the cards from a reflecting surface.

"Adding to the evidence for genuineness of the score is her behavior immediately preceding and during the placing of the cards. In a note which she had written to the experimenter while other children were working before her turn, she had said that this time she hoped to get twenty-five. The same hope had been repeated verbally before she began work. Just before her third run, on which she had the twenty-five hits, the experimenter shuffled the cards and handed them to her; she took them, but immediately handed them back, turned slightly away from the table, said 'Don't say anything, I am going to try something,' closed her eyes for a few seconds, turned back to the table; the experimenter meantime gave the cards another cut and handed them back to her; this time she placed the cards without cutting or shuffling them herself, all the time moving her lips as if she were speaking to herself. After she had finished, and before the cards were checked, she remarked, 'I was wishing all the time that I would get twenty-five.' This type of behavior had not, and has not since, been demonstrated by Lillian."

Discussion

The first question, of course, is whether the two exceptional scores of 23 and 25 may be regarded as produced by ESP. Only if this

is a reasonable position is it profitable to go on to the interesting questions which they provoke on that assumption. Obviously we have come to an affirmative answer to this question, or this paper would not have been written. A review of the reasons which led to this viewpoint is in order.

First of all, there is one especially solid bit of evidence that is not contained in the material quoted from MPR's report. Following the completion of the series of experiments mentioned above, MPR asked Lillian to come to her laboratory to take part in some tests with cards enclosed in sealed opaque envelopes. These results were incorporated in a report by JBR (1). In 25 runs which Lillian made on these enclosed card tests, she declined in score average well below the mean chance expectation of 5, averaging approximately 4. The total negative deviation of 24 in 25 runs gives a critical ratio of 2.4. Since it has been found that the enclosed card tests tend to induce negative deviations in the scoring of many subjects under certain conditions, it is not difficult to regard these results by Lillian as strongly suggestive of ESP performance.

Second, and perhaps of equal weight, is the fact that Lillian scored positively—that is, above mean chance expectation—on all four of the other methods—methods which are less open to criticism on the grounds of experimental precautions than is the OM method on which the two exceptional scores were made. On these four procedures the 80 runs totalled 454 hits, which, with its positive deviation of 54, gives a critical ratio of approximately 3.

Entirely apart, then, from the OM test procedure on which the scores of 23 and 25 were made, Lillian gave evidence of ESP ability. This is by no means proof that the two unusual scores are attributable to that capacity, but the knowledge that she demonstrated such capacity on other occasions and by other methods, places considerably less burden of *a priori* improbability upon the OM test conditions themselves.

But however inadequate we may regard the OM procedure because it permitted the subject to look at the backs of the cards and permitted her to handle them as well, it should be pointed out that with the exception of the two scores of 23 and 25, Lillian averaged actually below expectation on this type of test. If we deduct from her total of 137 hits for the twenty runs the deviation of 38 attributable

to those two runs, the remaining deviation is slightly negative. If, then, there was any advantage in scoring due to the inadequacy of the conditions of this test procedure, it must have been concentrated wholly in the two runs of 23 and 25 hits. Since these two runs appeared about the middle of the series of 20 for the OM procedure, this would appear unlikely. Furthermore, in the 8 runs that Lillian did on the OM procedure following the score of 25, the deviation was on the negative side, and there was no change in conditions, no additional precautions introduced to which we might look for possible explanation.

Taking these facts into consideration, along with the features reviewed in the quotation from MPR's record—the fact that the child did not look at the backs of the cards, that she was being carefully watched by the experimenter, that the table could not have reflected images from the cards, that the cards were especially selected for the purpose after the first score of 23 had been made—all together make the hypothesis that the two outstanding scores were effects of ESP the most reasonable one.

Added to this is the impression which MPR obtained from observation of the child's behavior through a reasonably close acquaintance of some months, which included, eventually, some visits to her home when she left the Refuge. This impression of the child was that she was entirely honest, and if any erroneous factors entered into the test results Lillian was unconscious of them. As the letter from Lillian to MPR given below indicates, a warm friendship had sprung up between the two. This may be a fact of some importance.

The Wright Refuge
Durham, N. C.
May 11, 1936

Dear Margaret,

I enjoy you coming. I am very glad you can come. We like to play cards with you. I hope you enjoy playing too. I am glad I got twenty-three of the cards right. But I want to get all of them right. I am going to try very hard to get all of them right. I think you and Peggy Price are very nice to us. I like to go to Duke University. I like you very much. I don't want to go home for I will miss you very much. I am going to try to come to Duke when I go home so I won't be lonesome.

With much love,
Lillian

On the assumption, then, that the two scores represented functioning of extra-sensory perception, it is important to consider what they show or even suggest with respect to the nature of ESP and the conditions under which it occurred. It is, of course, hazardous to attach final importance to instances as rare as this, at least so far as certain conclusions go. On the other hand, we have no hesitation in concluding that such occurrences, however rare, at least show certain possibilities respecting ESP.

First, a number of characteristics of ESP as it shows up in laboratory tests are confirmed. The unconsciousness of the ESP process and, because of that, the lack of control on the part of the subject so far as keeping it going or cutting it off is confirmed. The ability is turned on and turned off, as it were, without any awareness on the part of the subject as to when and how. The fact that Lillian could not repeat the performance after the second high score is good enough evidence that she did not know in the least how she did it in the first place.

The results, of course, give a certain stamp of approval to the general conditions that obtained; at least, these are shown to be possible producers of ESP scores. The use of children as subjects, the card tests, the matching procedure, the good personal relations between subject and experimenter, the use of small rewards, the informality of the test situation as contrasted to a formal laboratory routine—all these may be taken as possibly favoring the remarkable results, but certainly permitting them.

The points just mentioned could have been established quite as well with scores considerably lower than 23 and 25, but such scores as these do indicate unmistakably that there is a real possibility of perfect scoring and continuous ESP success. It is not difficult to believe that in the instance of the score of 25, Lillian possessed unerring ESP ability from the time she placed her first card right on through a period of probably 20 to 40 seconds until the last card was placed before its appropriate key-card. (In the case of Riess's subject when she scored 25, the period of the test covered 25 minutes.) This leaves the suggestion that it may not be unreasonable to expect that, if a measure of control may be obtained, the normally sporadic and spontaneous character of ESP performance may be changed to a continuous and stabilized order of action.

And it is precisely on this point of how to control the process of ESP that we think the performance of Lillian may be most helpful. The feature that seems to stand out in association with the two unusual runs is Lillian's determination to get them all right. It may have been her affection for MPR, as her letter would perhaps suggest, or it may have been her ambition to win the prize of fifty cents, or perhaps both together. At any rate, the intensity of her desire was so great that it impelled her into a kind of prayerlike ritual which, in effect, may have been a kind of spontaneous magic. She stated that she was "wishing all the time that I would get 25." May it not be that this little performance which she carried out before and during the test was her way of breaking through barriers of habits, associations, and rational processes that are normally much more available to conscious direction?

But if this were so, why did she not repeat the procedure and obtain another high score following the 25? One obvious answer is that motivationally she had reached a climax, and the great concentration of effort that had impelled her previously was automatically dispersed.

In this connection, we are reminded of the statement by the medium, Mrs. Eileen J. Garrett, whose ESP scores were significant in this Laboratory and were insignificant in ESP tests given her by Soal in London. She commented on the fact that she was under considerable pressure to score in this Laboratory, and felt no such pressure in England. In fact, she continued to improve up to the point at which she satisfied the statistical requirements for evidence of ESP, and then declined in scoring.

Probably every ESP experimenter will agree that subjects cannot obtain high scores in ESP tests simply by wanting very much to do so; otherwise, it would be necessary only to offer exceptionally high rewards in order to obtain high scores. But if there is a valuable suggestion in the performance of Lillian under discussion here, it would, we think, be that there is greater reason now to hope that if subjects can be made to desire very strongly to make high scores and can be guided at the same time into using a kind of spontaneous "magic" or individual mental device of nonrational character, necessarily peculiar to each person, we may be able to rescue ESP from its will-o'-the-wisp order of existence and subject it to a better con-

trolled performance. Such conditions as the individual may himself evoke to assist him in eluding the normal barriers to ESP performance may range all the way from small idiosyncrasies and eccentricities in the way he likes to lay his cards down or make his records, to automatism or deep trance. Certainly to the present writers, Lillian's case suggests the advisability of encouraging the use of such personal devices.

References

1. Rhine, J. B. ESP tests with enclosed cards. *J. Parapsychol.*, 1938, **2**, 199-216.
2. ———. Extra-Sensory Perception. Boston: Bruce Humphries, 1934.
3. ———. Some selected experiments in extra-sensory perception. *J. Abn. and Soc. Psychol.*, 1936, **31**, 216-228; also *J. Parapsychol.*, 1937, **1**, 70-80.
4. Riess, Bernard F. A case of high scores in card guessing at a distance. *J. Parapsychol.*, 1937, **3**, 79-84.
5. Warner, Lucien. A test case. *J. Parapsychol.*, 1937, **1**, 234-238.

THE ESP CONTROVERSY

By DOROTHY H. POPE and J. G. PRATT

WE HAD OCCASION to point out, in a former article (12) that a considerable portion of the material contained in this JOURNAL has been devoted to discussion, as distinguished from reports, of the research in extra-sensory perception and that some of these were in a critical vein. When the entire literature on the subject of parapsychology is taken into account, the articles dealing critically with the ESP work actually outnumber those which report original research. Since such a large part of the collective effort has been expended in this manner, the importance of judging the effect which the criticism has had upon the progress of the investigation is obvious.

There are two reasons why we wish to discuss this now: One is the fact that the literature of the controversy has by this time become sufficiently extensive that we should be able to generalize about it with considerable profit and so to derive some insight as to the present standing of the research and some guidance as to its future conduct. The second reason is that, writing as we are when the controversy has for the most part subsided, we have the advantage of being able to treat the critical phase of ESP literature in its entirety.

In view of the revolutionary character of the hypothesis of ESP for general psychology, it is not surprising that the literature of the controversy has, in this instance, been extensive. Indeed, it would have been surprising, after the experiments had received wide public attention, if considerable discussion had not developed. The history of science shows clearly that the claim for any unusual discovery will meet with objections to the degree that the principle discovered is unusual and the extent to which it receives general notice. Moreover, the ESP investigators, from the beginning, asked for criticism of their work, feeling that where the implications of the findings

would be so tremendous, every reasonable objection to the research should be examined and weighed.

Scope and Quality of the Criticism

There has, of course, been adverse criticism of extra-sensory perception ever since the beginning of research in this field more than sixty years ago. But the scope of our survey will be limited to the particularly active period of eight years which began in 1934 with the publication of the Duke experiments. The critical articles, reviews, and letters to editors which appeared in various newspapers, journals, books, and magazines during this time totaled approximately fifty-nine, and the frequency of their appearance, it will be noted, is roughly proportionate to the amount of public interest in the research. In 1935 and 1936, this interest was negligible and the articles in the press numbered only four and one, respectively, for those years. The boom in criticism came in 1937, along with increasing publicity. In the late fall of '36, Professor Ernest Hunter Wright of Columbia had written a popular review of the ESP experiments for *Harpers* (20), a condensation of which appeared in *Readers Digest* early in 1937. The Book-of-the-Month Club extended popular interest by its choice of Dr. J. B. Rhine's *New Frontiers of the Mind* (13) as their October, 1937, selection. The radio, capitalizing on current interest in ESP, carried programs over nation-wide hookups in 1937 and 1938, and most popular magazines and newspapers ran articles on the subject. Concurrently with all this publicity, the number of critical articles in the ESP literature increased to thirteen in 1937 and twenty-nine in 1938, nearly all of them representing new names whose owners were vigorous in their attack upon the research.

This same upsurge of interest in ESP was apparent at psychological meetings where the discussion of ESP had at first been heard only in the hallways and offstage, as it were. In 1937, the topic began to make its way into the programs, and in the spring of '38, several sectional meetings of the American Psychological Association listed papers on ESP, critical or otherwise. Finally, in the fall of that year, at Ohio State University, a special symposium on methods in ESP tests was arranged under the Association auspices. This event represented a climax in the controversy; and after it,

critical interest declined, not only at the psychological meetings, but in the mind of the public as well, a decline which was evidenced by a diminution in the number of critical articles to eight in 1939, four in 1940, and none at all, to our knowledge, since that time.

Mere amount of criticism, however, in terms of numbers of articles or pages affords no adequate conception of what the student of ESP wishes to know about the controversy. It is more important, among other things, to note who the principal critics were and what was the nature and quality of their contribution. The critics were, almost without exception, members of that broad classification known as "the psychological profession." This fact, of course, is quite as it should be, since the ultimate goal of parapsychology is the coordination of its findings with those of general psychology. That the controversy was, at this period in the history of the research, primarily the concern of psychologists may, therefore, be taken as an indication of real progress. In the period prior to 1934, never had half so much attention been accorded to the ESP research—not even to purely destructive criticism of it. But the preponderance of psychologists among the new critics indicated that the profession had at least been put on the defensive and a clear-cut issue had been drawn for the student of psychology.

The tone of the criticism in 1935 and 1936, which may be called quiet years, was moderately well tempered. But as public interest arose in '36, '37, and '38, there was noticeable an increasing irritation conveyed by explicit statement of condemnation, as exemplified, for instance, by Dr. Kellogg's designation of ESP as a "craze" (9). This emotional turn seems best interpreted as the professional psychologist's reaction to the frequent bombardment of questions regarding ESP which students and laymen leveled at him during the period when ESP was so frequently mentioned in the press and on the air. Given a natural doubt as a result of his training, the psychologist, unfamiliar with the experimental work as he doubtlessly was in most cases, must unavoidably have reacted with annoyance. When he set forth to clear the air of these disturbing claims, it was hardly to be expected that he could restrain his personal feelings without some difficulty.

Moreover, the claims of ESP genuinely conflict with the intellectual heritage of most psychologists. Since the 17th century,

if not indeed since ancient Greece, the social sciences have been predominantly influenced by the development of the more successful physical sciences. The result has been an increasing emphasis upon the more objectively checked processes of personal activity, the sensory and motor functions. Mental life beyond the senses is less subject to measurement in ultimately physical terms. In fact, the very erratic, spontaneous nature of extra-sensory perception sets it off as almost contrary to physical occurrences. Good systematic classification required that it be ruled out. It was easier to make the senses the only channels to cognition of the external world. This point of view has been symbolized by the classical assertion that there is "nothing in the intellect which was not first in the senses."

The facts of ESP therefore come into direct conflict with the prevailing scientific assumptions of centuries and this incompatibility in mode of thought, rather than any inadequacy of evidence, has made acceptance of ESP especially difficult for the psychologist. Ours is an age which has learned to see the world one way, and according to that one way, there can be no such thing as perception without the senses.

Review of the Criticism

The specific criticisms of the work are, of course, the more important consideration in this review. They fall most naturally into three general groups: (1) those concerned with the means of evaluating the ESP results with respect to chance—the statistics and mathematics used; (2) a general group dealing with the adequacy of the experimental procedures; and (3) those which may be characterized as pertaining to the logic of the interpretation of the ESP hypothesis, especially as it is related to the success or failure of different experimenters. These three topics are mentioned in an order which is itself of significance since it represents the general chronology of the criticism. Up through 1937, the objections to the research were chiefly concerned with the mathematics; through part of 1937 and in 1938, they were mainly directed at the experimental methods; and in the latter part of '38 and on into '39, they were concerned more especially with the interpretation of results. Each of these trends ran its course and each, as may be seen from what follows, reached a somewhat natural climax after which the issue apparently ceased to exist.

The mathematical phase of the criticism is represented, strangely enough, not by mathematicians but by psychologists. Not a single mathematician, in this period, raised his voice against the techniques of evaluation, and those criticisms coming from psychologists were sometimes obviously based on such uncertain ground as to require, in several instances, their subsequent withdrawal. The first critic of the mathematics was Dr. R. R. Willoughby, then of Clark University, who presented three papers on ESP in 1935 (17, 18, 19). His main point was that the method of evaluation in use was an improper one and he proposed an alternative. Later, however, he reversed his position on this point and still later, in 1940, conceded that the statistical issue was then irrelevant (15: ch. 8).

Dr. C. E. Kellogg, of McGill University, was another who, in his first article (8), suggested an alternative method of finding the probable error and withdrew it in a later discussion (10). His contention—that the standard deviation of the observed data should be used rather than that based on the theoretical binomial distribution—is not insisted upon in his last paper stating his critical position (15: ch. 8) and it seems safe to infer that he has given up that point also. At any rate, he recognizes clearly that portions of the evidence differ significantly from chance. Obviously, from the standpoint of whether or not ESP occurs, any further discussion would be merely academic.

In any case, a satisfactory answer was provided, for the question which Kellogg raised, by Greenwood's empirical chance study of card matchings (5). This study consisted of 500,000 matchings of ESP cards against actual calls made by subjects in earlier tests with which they were not intended to be paired. He found the frequency of correct correspondences to be very close to binomial expectation; that is, the calls represented a chance series. On the basis of this study, Greenwood was able to show that the methods actually in use in the ESP research—methods based on the binomial expectation—were more nearly applicable than those proposed by Kellogg.

The question arose, too, regarding the accuracy of the probability of one fifth attributed to each trial in the ESP tests. In other words, it was doubted whether five hits per twenty-five was the correct expectation on the theory of chance. Mathematicians and ESP research workers, however, were able to show by empirical control

series and by actual mathematical proof that this objection to the mathematics was completely groundless.

These by no means exhaust the range of the mathematical questions and criticisms which were raised; but there was only one other which was sufficiently relevant to deserve mention. This one—a suggested correction in the method of computing the standard deviation—came principally as a result of Kellogg's criticism and led to the use of an alternative technique by ESP workers for a time. Later on, the large empirical chance series by Greenwood, which was mentioned above, settled the issue in favor of the method first in use, and this correction of two percent of the standard deviation was discontinued.

We have remarked that those who criticized the statistics of ESP were not mathematicians, and during the mathematical phase of the period of controversy, this was literally true. Recently, however, there has been an exception, one that is entirely "out of line" with the other members of his profession who have generally approved the methods in use by ESP workers. It was some three years after the mathematical issue had been laid to rest by the mathematicians themselves that Dr. Willy K. Feller (4) attempted to dispose of the case for ESP on the ground that the cards were inadequately shuffled, the data improperly selected, and the experiments spuriously significant because they were terminated at favorable stopping points. However, Greenwood and Stuart (6) were, without difficulty, able to show from the research reports already published that these criticisms could not account for the ESP results nor invalidate the conclusions. Feller has not resumed the discussion since that time.

In assessing the value of the mathematical controversy, it is fair to say that while mathematicians did not openly criticize the ESP work, they did, in response to requests for assistance and advice by ESP workers, contribute greatly to the integrity of the work by offering suggestions for proper procedures and giving their assistance when mathematical difficulties arose. In fact, the ESP controversy brought forth a quantity of original statistical research. When the subject of ESP statistics finally was raised in 1937 at the meeting of the Institute of Mathematical Statistics, a now familiar press release was issued by the president of that body approving, in gen-

eral, the validity of the statistical analyses of the ESP work and ending with the statement: "If the Rhine investigation is to be fairly attacked, it must be on other than mathematical grounds" (1). This release and the publication of several other explanatory articles marked the abrupt cessation of mathematical controversy in relation to ESP. Appropriately, Professor E. V. Huntington, in the most outstanding of these articles (7), inquired: "If mathematics has successfully disposed of the hypothesis of chance, what has psychology to say about the hypothesis of ESP?"

* * * *

The critics had already begun to scrutinize the experimental adequacy of the ESP methods even before the authoritative ruling of the Institute of Mathematical Statistics was made public, and in 1938, the year following that statement, there was a blast of criticism on this new note. Most of it had to do with a total misunderstanding by the critics of the use of the commercial ESP cards in the techniques employed in the research. The production of the ESP cards had, in a general way, been supervised from the Duke Laboratory; but in spite of the warnings and instructions given, the manufacturer was not able to avoid a certain amount of warping of the cards which, while it did not show up in proof, appeared some time after manufacture. While this warping of the cards was not at all obvious to the average person, the symbols could be read from the backs of the cards if they were held at a certain angle to the source of light.

The decision had to be made by the members of the laboratory staff as to whether the entire stock should be discarded or whether it should be released, with appropriate instructions, of course, concerning the use of the cards for experimental purposes. It was decided that the latter course was advisable, if not unavoidable, since a reprinting could not be obtained. For the layman's purposes the cards were satisfactory for personal testing and illustrative material; only the person who was seeking to deceive was likely to avail himself of the possible cues. And for scientific testing, no cards could be considered "perfect" and none could be reliably used without screening, anyway. An announcement in the next number of the JOURNAL OF PARAPSYCHOLOGY suggesting the use of opaque screens with the ESP cards afforded sufficient warning. For that matter, experienced investigators had long since ceased to trust to

any cards, however free from defects they might seem to be at the beginning of a test, and no research had been reported from the laboratory in which the primary conclusions had been based on tests made with unscreened cards.

But in spite of the notice in the JOURNAL and with utter disregard for the emphasis upon screening of cards in the ESP reports, Dr. S. H. Britt read a paper at the spring meeting of the eastern branch of the American Psychological Association in which he made a sensational demonstration of how the commercial ESP cards could be "read" from the backs. Other critics—Kennedy, Skinner, Wolfle, Gulliksen—joined in, all completely ignoring the fact that the cards, however faulty, could not give sensory cues *if they were concealed behind opaque screens.* This was not the only evidence abroad that the actual reports had not been carefully read.

So much was made of these defective cards and so much attention given to ESP research in 1938 that the A. P. A. arranged for a symposium on experimental methods of testing ESP (3). It is perhaps indicative of the temper of the times that Kennedy, then one of the most active critics of the ESP work, was appointed the chairman. His own paper—suggesting that recording errors would account for the ESP results—was so completely and adequately answered by Dr. Gardner Murphy, who followed him, that this brand of criticism has not been raised since that distinctive occasion. Moreover, Gulliksen, who led the attack on the issue of experimental inadequacy, which had been the main current charge against ESP for a year, agreed on the floor that the experimental methods as described by Rhine were, indeed, acceptable to him.

Thus ended the second phase of the criticism, the problem of experimental methods. There was little or no questioning, by this time, of the mathematics in use and the A. P. A. symposium more or less closed the issues in existence up to that time. It was manifestly a turning point in the ESP controversy and few critical papers have been written since then. It had the effect of clarifying the issues and it accomplished a great deal toward the general understanding of the character of the ESP research.

* * * *

The third period of criticism overlaps the second as the second overlapped the first. It was in the summer of 1938 that Dr.

Clarence Leuba (11) voiced his doubts as to the conclusiveness of the ESP results on the grounds, first, that an experimenter could take advantage of a convenient run of luck by "optional stopping"; i.e., by stopping his experiment when results had been favorable; and second, that due consideration was not given to the unsuccessful tests by experimenters who had tried to get positive results and had failed.

The optional stopping topic was current in the discussions of ESP throughout 1939 and was one of the main issues in the ESP symposium held by the Southern Society for Philosophy and Psychology that year. Indeed, it is one of the few published criticisms which actually led to an alteration of methodology, even though, it should be said, such alteration was not essential for the establishment of the ESP hypothesis. Briefly, the question of optional stopping was solved, so far as the present period of controversy is concerned, not so much by argument as by methodological adjustments. Experiments were simply given, in advance, an estimated range which did not allow for the option in question. However, the previous policy of *publishing all the investigations made to date* under the specified conditions was itself, for the time, an effective answer to the question raised, even though technical improvement was possible. Since a complete report cannot be criticized for selection, this question, too, was only an academic one and did not concern the main issue of the ESP hypothesis. All that remained for the critics to say was what Kellogg finally did say, in effect, in his 1940 critique (15: ch. 8): "Keep on going under present test conditions and you will probably find that you simply will not get any more evidence of ESP." Every issue of the JOURNAL OF PARAPSYCHOLOGY since that time has been testimony to the failure of his prediction.

Of far more concern among psychologists and, through them, among critical lay-readers, was the other charge, that not enough importance was attached to the "negative series" the results of which did not show ESP. By some odd type of reasoning, certainly not mathematical, it seems to have been felt that a failure on the part of one psychologist to obtain evidence of ESP when he administered his tests—whatever may have been the conditions or methods he used—canceled out the favorable findings of another. In this regard, people seem to have been misled into comparing ESP

with the physical sciences in which, if one investigator does not confirm another's work, there is something obviously wrong somewhere. It is forgotten that human beings are not like inorganic bodies and substances; they are exceedingly variable and subject to the most subtle influences. The conditions of the experiment are very important, and individual differences sometimes extreme. Consequently, one psychologist's failure may actually mean nothing regarding another's success.

The most vigorous critic in this aspect of the research was an Englishman—not a psychologist, but a college teacher of mathematics—Mr. S. G. Soal, who had himself conducted a long series of ESP tests from which he had not obtained a significant positive deviation. Contrasting his results with those of American investigators, he threw out strong implications against the propriety of reaching favorable conclusions from work that did not agree with his own. Since 1940, however, this critic has been led to reverse his stand (16), and that, strangely enough, by his own discoveries which are in themselves among the most interesting of recent years. At the suggestion of Whately Carington, he looked for displacement effects (that is, ESP hits on neighboring targets rather than on the intended targets) and he has turned up highly significant evidence of ESP. This evidence has apparently continued to accumulate even under the difficult conditions induced by the war. Hence, the most conspicuous "failure" has turned into what is perhaps at present the most remarkable "success." Little is heard or probably will be heard of this third stock argument against ESP research, for this one instance goes far to remind the critics in general that determination of success and failure may lie even within the experimenter himself, let alone those other variables—subjects, conditions, and methods of analysis. "Failure," then, may be an indiscriminate term.

From a consideration of these three phases of the controversial era of ESP, it would appear that events, more than arguments, turned the tide of criticism—events which have brought into focus, or at least brought to the attention of the audience concerned, the arguments or facts which would have otherwise been inadequate to stem the criticism. Specifically, it was the official pronouncement by the Institute of Mathematical Statistics, the A. P. A. symposium,

and Soal's discovery of displacement in his data that have marked the turning points, if not, indeed, constituted them.

This is not to say that there is general acknowledgment of the finality of the findings or agreement as to the interpretation of the results. It is safe to say, however, that there is a fairly widespread recognition of the scientific character of the research and an increased respect for the way it is being done. When, in the fall of 1939, the seven leading psychological critics—Willoughby, Kellogg, Wolfle, Gulliksen, Kennedy, Lemmon, and Thouless—were invited to produce their criticisms for incorporation into the book, *Extra-Sensory Perception After Sixty Years* (15), a marked change in attitude from that of 1937-38 appeared, one that was obviously more thoughtful and restrained than that which some of the same writers had shown in their earlier critiques. In fact, it should be said (particularly because objective criteria on such matters are so few) that only three of the seven accepted the invitation. This, it must be remembered, was only a year or two after the time when those who now declined had been very actively engaged in vigorous criticism. Of the three who accepted the invitation, Dr. R. H. Thouless, of Cambridge University, never a completely destructive critic in the first place, sounded a wholly constructive and favorable note. Dr. V. W. Lemmon, of Washington University, also customarily temperate in his criticism, was restrained and relatively neutral. Only Dr. Kellogg continued to reject the ESP research without compromise; but the force of his criticism had been greatly moderated and was by no means difficult to cope with. In a word, the response of the seven critics tells more than any other single event of the end to which ESP criticism has finally come: mainly silence, with some shift toward favorable, or at least neutral, attitudes and an element of restrained but persistent die-hard rejection.

Effect of the Criticism on the Research

Turning from these more specific features of the ESP controversy to the matter of its general effects and significance, we find ourselves on more difficult ground. Undoubtedly the past eight years have witnessed marked advances in the methodology of parapsychological research, as regards both the statistical evaluation of the results and the experimental methods. The question which we wish finally to

raise is this: To what extent have these advances come about as a result of the published criticism?

An answer to this question cannot be given without our first calling attention to the fact that the best type of criticism of the research has generally not been published; consequently, it never became a part of the public discussion, that exchange of articles which have been grouped under the general term, the "ESP controversy." In 1938, when the discussion of the ESP results was at its height, the editors of the JOURNAL OF PARAPSYCHOLOGY (2) pointed out that the best criticisms of the experiments had been offered by fellow investigators. Next, as a class, in fruitfulness of their suggestions came the active students of the research and the intimate colleagues of the investigators who offered their constructive criticisms of the work as it progressed. Much of the discussion within and between these two groups took place either in conversation or through private correspondence so that there was nothing for the scientific world to see except the *effects* produced on the research itself. The general result was such real progress toward complete safeguarding of the tests that the first publication from Duke University contained some of the soundest evidence of ESP which has yet been produced. Many of the hostile critics who later published attacks upon the experiments raised only such objections as could be effectively answered by the evidence already at hand. It was the rule rather than the exception that the objections of the critics were already being dealt with even before they appeared.

It seems to us that this editorial commentary is still a good evaluation of the role of the critics, and inasmuch as we have dealt here exclusively with the adverse attacks upon the research which were featured in the actual controversy and which had been anticipated in the natural progress of the experiments, there is not a great deal left to credit to this class of critic. Moreover, those objections which had some basis in fact were presented in a manner calculated to obstruct and weaken the research, and the resolution of the problem usually fell upon someone more constructively disposed toward the field of parapsychology. But the issues were eventually settled; and, in fairness, the critics may properly be credited with having helped to produce certain refinements of methodology which the investigators themselves had either overlooked

or had not regarded as essential. The provisions against optional stopping and for independent recording are in this classification. They are precautions which are now taken in most ESP research but not expected of other fields of experimentation. Whether they are requisite for all ESP work is still perhaps a debatable issue.

To close the account, it is equally necessary to total up the losses which parapsychology has encountered through the ESP controversy. The first item is that of the time and effort which have gone into answering critics and the attempt to anticipate what they might later have to say. That this time and effort have been onerous is self-evident from the amount of the ESP literature which has dealt with or touched upon the controversy. It is impossible to say exactly what difference it would have made for the history of parapsychology if the investigators had been able to use the many hours so spent in further research. We can only say with certainty that the controversy has been a great drain upon the time and energies of the experimenters.

It has also caused the experimenters to change their general approach to ESP testing. The change has most noticeably affected the psychological atmosphere of the tests and the experimental freedom used by the investigators. The term "psychological atmosphere" is used to refer to those delicate social factors which are difficult to describe objectively but which research workers have long recognized as essential to the successful demonstration of ESP. They inhere in the condition of rapport between the investigator and his subject. A spontaneity of interest in the test and its outcome has been the essential condition most commonly specified. There can be no doubt that as investigators have more and more planned their tests in the light of what the critics would think, they have so formalized the procedures that it has been increasingly difficult, if not impossible, to arouse a favorable type of interest in the tests.

This sacrifice of experimental freedom to the critics was made in an effort to remove even the appearance of evil from the research. Because there have been objections to tests which were made with the cards in sight of the subjects, the ESP researchers have gone to the other extreme of *never* doing tests in this fashion. Because the use of exploratory tests led to the charge of keeping the good results and throwing away the bad ones, preliminary trials for the purpose

of selecting suitable subjects and for "warming up" have been abandoned almost altogether. Because unconsciously motivated errors were attributed to the investigators, experiments were planned with a careful division of labor between two observers so that errors short of collusion were impossible. By such changes as these, investigators made concessions which silenced the critics but which, at the same time, may have served to hamstring their own activities since the occurrence of ESP is subject to certain influences which are known to be easily affected by the experimental conditions.

In any event, the first step is to secure the effect to be studied. It is like the old recipe for cooking 'possum: first you catch the 'possum. Accordingly, it may well be that in ESP testing the informality and spontaneity which accompany unstudied preliminary conditions is the surer way of leading up to an adequately safeguarded, crucial demonstration of the phenomenon. To many of us, the Duke experiments appear to suggest that this is the case, for it was out of such informal beginnings as those with Pearce, to cite a familiar example, that the Pearce-Pratt series (14) and other rigorously safe-guarded and successful projects emerged.

Another large item on the debit side is the effect which the controversy has had upon professional groups outside the field of psychology and upon the general public. Individual psychologists are, after all, in a fair position to form their own opinions about a body of evidence which is so vitally related to their branch of science. Other professional groups would be much more likely, however, to be influenced against the conclusion that ESP occurs simply because there are professional psychologists who, at least at one time, emphatically asserted that the evidence was unsound. The non-psychologist may not be aware of the relatively unsettled state of the whole science of psychology. When he sees or hears a man who has been trained in that branch of study dispose of the entire case for ESP, he is likely to agree out of deference for the critic's position. There are known instances in which this has been so and we can only guess how many more people have been influenced to take their opinions of the case for ESP from unjustified attacks on the experiments by psychologists who have published their criticisms with an air of finality that has sometimes approached the ridiculous.

By way of final account, let it be said that the investigators will,

as ever, continue to welcome criticism of their results and will doubtless, in some instances, deserve it. But they may be excused for holding to the hope that future criticism will come closer to dealing with real issues in the research than it has in the past. The best prospect for the progress of the research would now seem to require that the main points of this past controversy—many of which need never have been raised—be regarded as settled for good. Thus will investigators and critics alike be free to use their time most constructively in regard to the real issues of the evidence and the great problems of its interpretation that lie ahead.

References

1. Camp, B. H. [Statement under "Notes"], *J. Parapsychol.*, 1937, **1**, 305.
2. Editorial. *J. Parapsychol.*, 1938, **2**, 77-83.
3. ESP symposium at the A. P. A. *J. Parapsychol.*, 1938, **2**, 247-272.
4. Feller, Willy K. Statistical aspects of extra-sensory perception. *J. Parapsychol.*, 1940, **4**, 271-298.
5. Greenwood, J. A. Analysis of a large chance control series of ESP data. *J. Parapsychol.*, 1938, **2**, 222-230.
6. Greenwood, J. A. and Stuart, C. E. A review of Dr. Feller's critique. *J. Parapsychol.*, 1940, **4**, 299-319.
7. Huntington, E. V. Is it chance or ESP? *Am. Scholar*, 1938, **7**, 201-210.
8. Kellogg, C. E. Dr. J. B. Rhine and extra-sensory perception. *J. Abn. and Soc. Psychol.*, 1936, **31**, 190-193.
9. ———. New evidence (?) for extra-sensory perception. *Sci. Monthly*, 1937, **45**, 331-341.
10. ———. The problems of matching and sampling in the study of extra-sensory perception. *J. Abn. and Soc. Psychol.*, 1937, **32**, 462-479.
11. Leuba, Clarence. An experiment to test the role of chance in ESP research. *J. Parapsychol.*, 1938, **2**, 217-221.
12. Pope, Dorothy H. and Pratt, J. G. Five years of the Journal of Parapsychology. *J. Parapsychol.*, 1942, **6**, 5-19.
13. Rhine, J. B. New Frontiers of the Mind. New York: Farrar and Rinehart, 1937.
14. ———. Some selected experiments in extra-sensory perception. *J. Abn. and Soc. Psychol.*, 1936, **31**, 216-228. [See also *J. Parapsychol.*, 1937, **1**, 70-80.]

15. RHINE, J. B., PRATT, J. G., STUART, C. E., SMITH, B. M., WITH GREENWOOD, J. A. Extra-Sensory Perception After Sixty Years. New York: Holt, 1940.
16. SOAL, S. G. Fresh light on card-guessing—some new effects. *Proc. S. P. R.* (London), 1940, **46**, 152-198.
17. WILLOUGHBY, RAYMOND R. Critical comment: the use of the probable error in evaluating clairvoyance. *Char. and Personality*, 1935, **4**, 79-80.
18. ———. A critique of Rhine's "Extra-Sensory Perception." *J. Abn. and Soc. Psychol.*, 1935, **30**, 199-207.
19. ———. Prerequisites for a clairvoyant hypothesis. *J. App. Psychol.*, 1935, **14**, 543-550.
20. WRIGHT, ERNEST HUNTER. The case for telepathy. *Harpers*, November and December, 1936.

A CONFIRMATORY STUDY OF SALIENCE IN PRECOGNITION TESTS

By BETTY M. HUMPHREY and J. B. RHINE
Parapsychology Laboratory, Duke University

ABSTRACT: In one respect, this paper is a repetition of an earlier report made by one of us, in which was presented evidence of precognition obtained by means of salience ratios measuring certain peculiarities of hit-patterns within the run. This report confirms the earlier findings both as to terminal salience and precognition within the run. In another respect, it confirms and goes beyond the previous investigation. In that work there was a comparison between two-day and ten-day precognition tests, with the subject knowing in both cases approximately when the check-up would be made. In the present case, the subject was told only that his test would be divided into two parts, one half to be checked immediately, and the other half to be checked after ten days. The division was made on a random basis by the investigators after the calls were made, and the subject had no way of knowing in advance which runs would be in either section. Thus the procedure represents a precognition test which, so far as the subject was concerned, had an uncertain time period.

The results of the present experiment, like those of the three preceding reports on salience from this laboratory, show significant evidence of salience. This effect is found, as in the earlier instances, primarily in a significant covariation-CR between the salience ratios of the run and those of the five segments of the run. Under the circumstances of the test, such extra-chance relations are considered evidence of precognition.

Again, as in two of the three earlier studies, it is the larger groupings—Divisions and Sections—which show the significant interrelations between salience ratios, while the Subseries show only insignificant covariation-CR's. Likewise, too, this extra-chance relation

is supported, and significantly so, by covariation between the salience ratios of the two main Divisions, representing the child and adult subjects. Also, a U-shaped curve is given by the distribution of the pooled deviations for the five segments of the run, as was true of the preceding work of similar nature.

A notable feature of the experiment is that the more significant part of the results is contributed by the Section with delayed checking. This was found also to be the case in the preceding report. The extra-chance effect is mainly due to the child subjects. Their work is significant alone.

Introduction

As the title of this paper implies, we are concerned with a relation between salience and precognition. More specifically, the methods of measuring salience are used as the basis for determining whether or not precognition occurs in the experiments reported. By salience is meant the relation of the rate of success in the ends of a run (or segment thereof) to that of the middle segments (or trials of the segment). And by precognition we mean knowledge of uninferable future events.

In certain respects, the work described in this paper confirms the findings of an earlier paper on salience in precognition tests (18), and it also confirms the main results of two still earlier papers on salience (17, 19), dealing with the DT procedure instead of precognition. In order to make the extent of confirmation more clear, it will be necessary to review in brief the main results of some of these earlier papers.

In order to make reference to the various articles less confusing, we shall adopt the Series numbers used in the paper just preceding this one. All the work reported in the first terminal salience paper (17) is referred to as Series I. This involves only DT tests. (The second paper on salience had to do with a transoceanic series and was not part of the series of experiments conducted specifically for the purpose of finding and varying the amount of salience in ESP tests. Therefore, it will be excluded from further discussion on the results of the earlier salience papers.) The third article (18) reported

Series II and III, both of which are precognition experiments. The work reported in this paper will be called Series IV, and, as stated, consists of precognition tests, or PDT.

From the Viewpoint of Precognition

In the report on Series II and III, a review was given of the reasons for conducting those Series under the conditions that were followed. Briefly, the 2,302 runs of those Series were done with the subjects writing down their predictions of the order of decks of cards which were subsequently to be randomized by machines and cut on the basis of temperature readings of a specified date. In one Series, there was a two-day lapse between calling and the final check-up; in the other Series, ten days intervened. Up to the time when Series II was begun, mechanical shuffling had been the method used for randomizing the cards (7, 15, 23). The introduction of the temperature readings as a chance factor in determining the order of the cards, then, constituted an improvement. It was believed, or at least supposed, that this might offer a better basis of randomization than any kind of mechanical selection that had been used up to that time.

The fact that the results as a whole for these two precognition Series (II and III) were significant as measured by the covariation-CR, and the fact that the ten-day Series gave a considerably higher covariation-CR than did the two-day Series, would appear to justify the conclusion that the longer time period was not a limiting factor or condition.

Hutchinson (7) had found, in a one-day and ten-day comparison in precognition tests, that only the one-day gave extra-chance results. The explanation for the adverse comparison of the ten-day to the one-day tests may lie in the fact, however, that Hutchinson's subjects did both the one-day and the ten-day at the same session and were given their one-day scores and the rewards as soon as the checking was done, thus placing the ten-day tests at a considerable disadvantage so far as the subjects' interest was concerned. At any rate, the results of Series III in the salience experiment show that ten-day precognition is possible.

The question was raised by these studies as to whether the subject in any way takes note of the different lapses of time between

prediction and check-up, whether any real cognizance of the intervening period is important. The question might be put as follows: Does it make any difference to the subject whether it is a one-day or ten-day test? Or still more generally speaking: Does it make any difference to the subject when or how the check-up is made? The research we are now reporting is intended as a preliminary step toward answering this question.

From the Viewpoint of Salience

The studies on salience began, as the reviews in the earlier papers on the subject indicate, with the appearance of U-shaped curves of hit-distribution in the run plotted in five-point totals, each point representing five trials. These U-shaped curves were reported first in 1934 for DT test results (16), and were confirmed later by further experimentation (2, 3, 12). There were exceptions (8, 11); but even with the exceptions it was generally true that the ends of the runs tended to show greater deviation from expectation (in one direction or the other) than did the middle sections. This standing-out of the ends, or "terminal salience" as it is now called, appeared to be characteristic of the DT tests.

In the first article on terminal salience, a U-shaped curve was reported, though in this case it was an inverted one. In view of the fact that the total deviation was significantly negative, it seemed justifiable to regard this type of curve as entirely in line with the normal U-shaped curves obtained in earlier positive series. It is especially in line with the inverted U-shaped curve found by Pegram (12) in her Low Aim Series, which gave a negative deviation. The results of Series II and III combined give a normal U-shaped curve that is quite as striking as any that has been obtained.

A measure was introduced for the evaluation of salience (5). It is called the salience ratio, or SR. By means of the SR, the relations may be discovered and compared among the various subdivisions of the experimental data. It is in these interrelations that the significant features of the results from Series I to Series IV inclusive have been found.

In Series I, the ordinary DT run was broken up into segments of five trials each by the introduction of interruptions at the proper intervals. As measured by the salience ratio, salience was found not

only for the run as a whole, but for these segments themselves. In other words, as miniature runs they showed the same effect of salience as the run itself. Moreover, it was found that a significant relation obtained between the salience ratios of the segment and those of the run when these ratios were obtained from sufficiently large subdivisions of the experimental results. This did not hold true, however, if the smallest subdivisions (Subseries) were used as the basis for the measurement of interrelations. In other words, the relation between the SSR (or segmental salience ratio) and the RSR (run salience ratio) as measured by covariation statistics indicated clearly that an extra-chance relationship was present.

Although this SSR-RSR covariation was the principal basis for concluding that the results were not explainable by chance and that they constituted evidence of precognition, other covariation measurements did support this finding. Both in Series I and in Series II and III combined, there was some relation shown to exist between the SR's of the Child Division of the Series, and those of the Adult Division, though in no case was this relation quite as marked as that of the SSR-RSR relation itself.

Since the results of Series II and III confirmed those of the first study, Series I, as regards the significant evidence of terminal salience, it seemed justifiable to hope that this phenomenon of salience might be relatively repeatable. This hope was further warranted by the somewhat lengthy history of U-curves which preceded the more general analysis of salience. It was, then, with an optimism born of this reflection, that further work was undertaken, with a view to using the salience effect, if it could be obtained to a significant degree, as a measure of extra-sensory perception, as it had been in Series I, II, and III.

More elaborate statistical studies, either under way or in contemplation, are intended to throw further light on the nature of the salience effect that has been obtained. For the purpose of the present study, however, little has been intended, so far as salience is concerned, beyond using that phenomenon as a secondary means of pushing our knowledge further ahead on the primary question of whether uninferable future events can be predicted; and, if so, under what conditions.

Description of the Experiment

The Plan

As with Series I, II, and III, Series IV consisted of as near to 1,000 runs as it was convenient to come. But while, in the first three Series, the last Subseries somewhat exceeded the even number of runs, it was possible to end this one exactly on the 1,000 mark. Again, as with the preceding Series, the experiment was done in batches, the exact size of which could not conveniently be predetermined, since it depended upon the number of subjects involved and the interest of each subject. These batches make up the smallest subdivisions, or Subseries, as described below. No option was exercised in the control of the size of these Subseries, in the light of their success, since each Subseries was closed before any checking or evaluation was carried out.

Series I, II, and III were conducted during the spring semester of 1941; Series IV was done two months later, during August and September of the same year.

At the time the experiment of Series IV was begun, the results of Series I, II, and III had been fairly completely analyzed statistically, and it was possible to use the results to advantage in planning the next step. This had not been the case in the preceding Series. We decided to use the broken-run technique, which had been used to some extent in Series II and III. This divided the run into five segments by means of the structure of the printed record sheet and not by interposing an interruption test as in several previous Subseries. (See details below.) The results from the preceding paper also warranted a further attack on the comparison of precognition over a short, as compared to a long, period. Since the ten-day precognition test had succeeded, according to the measures used, quite as well as the earlier two-day work, it was decided to go on with that comparison, with the introduction of an added feature. It seemed highly important to attempt to find out what role the length of time between prediction and checking really plays. Toward this end, in Series IV the subject himself was kept from knowing by any normal means whether he was doing a given run for immediate or for a delayed check-up. Half the results would be checked as soon after they were made as the conditions permitted; the other

half would be checked ten days later. Presumably, different results in such an experiment would make considerable difference as to the hypothesis of precognition, when the time arrived for the consideration of that phase of the research.

As in the earlier Series, the separation of subjects into two groups on the basis of age was a part of the experimental plan. Accordingly, the headings "Child Division" and "Adult Division" will be continued in this presentation. (See Figure 1 for a diagram of the experiment.)

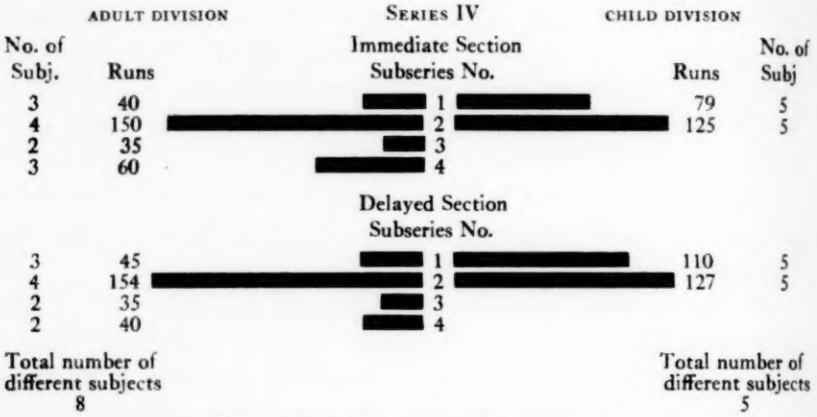

FIG. 1.—DIAGRAM OF THE EXPERIMENT

In the diagram it will be noted that the immediate and delayed subdivisions are indicated as Sections. Under the Adult Division there are four Subseries in each Section, Immediate and Delayed, while under the Child Division there are two in each Section. The diagram represents the approximate size of the Subseries, and indicates in terms of figures the actual number of runs in each. The number of subjects participating is likewise given.

As stated, the Divisions represent the subjects separated according to age; the Sections represent the test conditions which, in this case, mean simply the differences in time before checking. The Subseries represent the number of runs rounded up for each check-up, and include all the records that were available at the time of the check-up. The number of runs varied according to the number of subjects available, or to the interest taken in the test. They ranged from 35 to 154 runs.

Subjects

The subjects who participated in the tests of Series IV had nearly all taken part in Series I, II, and III. However, Series IV was conducted at Lake Junaluska, North Carolina, where a part of the staff of this Laboratory spent the summer. Only a small part, therefore, of the group of subjects who participated in the earlier Series was available for this one. There were five children and eight adults.

Again, as in the case of the former Series, there will be no analysis on the basis of individual performance and no further description of subjects. This aspect of the research will be the subject of a later report after the necessary statistical work has been completed.

General Conditions and Attitudes

Both children and adults worked individually in filling out their record sheets, but an effort was made to avoid exerting any pressure upon the subjects to undertake the tests for any reason except their own desire. Thus, most of the tests conducted with the children were ones in which the subjects had asked for the record sheets. Presumably, they did this partly because of an interest in the rewards; but to the experimenters it appeared that there was also a certain pastime or entertainment value in filling out the blanks. At least, the children were likely to ask for record sheets when they had nothing else to do. On the other hand, the adults almost certainly had no interest in the monetary value of the task itself. There was, however, some intellectual interest in the possible outcome, though in at least two instances adults who had participated in Series I, II, and III did not turn in blanks that were given them to be filled out. One other filled out only ten runs. We can be certain that the children were interested in the task, or they would not have participated. In the case of the adults, we are doubtful if the interest was very great.

Materials and Apparatus

The regular ESP cards were used, five decks for each batch or Subseries. The shuffling of these cards was done by the Merritt Shuffling Machine.[1] The maximal and minimal temperature read-

[1] See Volume 2 of the JOURNAL, 1938, pp. 231-232.

ings used for the cutting of the cards were taken from the Asheville *Citizen*. The special record sheets used in Series IV have been used in preceding work. Each contains space for one run; there are five columns of five spaces each, twenty-five in all. Each space is divided into two parts, the left for the subject's prediction, and the right for the target symbol which is to be selected later by chance.

Procedure

The subjects were supplied with the record sheets and were asked to attempt to identify the symbols which would later be hectographed in the right-hand spaces of each column. They were instructed to record their choices of symbols in the left-hand column opposite and to fill out the spaces in consecutive order.

Each subject was allowed to do as many runs as he wished, to do them at whatever rate suited him best, and to turn them in to the experimenter whenever he was ready. The record sheets were collected and held until fifty or more were assembled. When that point was reached, the rule was followed that the checking would be made on the next day, using the figures for the extremes of temperature for the day preceding, as published in the morning paper. These conditions governing the time of checking naturally allowed for much variation in the length of time between the filling-out of the record sheets and their first checking. This varied from one day to more than a week, in some instances. For the most part, however, the time interval was one and two days. All of the subjects but one were in the immediate community. The one exception sent in his records by mail.

Regarding the checking procedure, subjects were told that one half of their total number of record sheets (every other five) would be checked on the day following receipt of as many as fifty altogether; and that the other half would be checked ten days later. It was explained that whether the odd fives or the even fives were to be checked would be determined by chance. The record sheets of each subject were to be numbered consecutively. Thus ten or more runs from each subject were needed for this division into immediate and delayed check sections.

The child subjects and one adult were told that they would

receive small monetary rewards (for a score of eight, five cents; score of nine, ten cents; score of ten, twenty cents; score of eleven, thirty-five cents; and so on, with double rewards for scores in the Delayed Sections).

After all the record sheets were received by the experimenters, they were turned over to JBR and divided first into Adult and Child Divisions and then into Sections, thus making four subdivisions. Which subdivisions were to have immediate and which delayed checking depended on whether the total of the maximal and minimal temperature figures was an even or an odd number. If odd, the first, third, fifth—or, in general, the odd—fives of numbered record sheets were checked immediately, and the even fives placed in the Delayed Section. BMH obtained from the morning paper of the day on which the checking was done the figures giving the temperature extremes of the day preceding. The smallest digit of these two numbers, except zero, was taken as the number of revolutions which each of the five ESP decks and a numbered deck were to be given in the mechanical shuffler. The total obtained by adding all the digits of the temperature figures gave a number which was used to cut a deck of twenty-five numbered cards. The number turned up on this cut was then used in cutting each ESP deck after it had been taken from the shuffling machine. If the addition of the temperature digits exceeded twenty-five, then twenty-five was subtracted and the balance used to cut the numbered deck. The five decks were then recorded, and at that point JBR turned over the record sheets to BMH, who hectographed the card records on the appropriate sheets; that is, the deck designated as number one was hectographed on the first one of each block of five record sheets. Thus, if there were, for example, one hundred record sheets, each card record would be used twenty times, and several of these twenty record sheets might have been filled out by the same subject if he had more than one block of five records in that Subseries. This condition will be discussed in a later section of the paper.

When the hectographing was completed, BMH checked the records, and, in order to make the checking independent, left no record of the first checking on the sheet. Miss Jean Fischer, an undergraduate assistant, and Mr. Roy Upshaw, of the secretarial staff, made a second check-up, marked the hits, and recorded the

scores on the record sheets. At the completion of the check-up, the checkers compared scores; and if a difference was found, the run concerned was re-checked by both. Miss Fischer and Mr. William Birge, a graduate assistant, finally gave the results a third check.

At the completion of the second check, and after the total number of hits for each position in the run was obtained, the assistants made up the deviation-distributions for the five segments of the run and the five trials of the segment—two distributions for each Subseries. These distributions were checked against the score totals themselves, and, failing agreement, were re-checked. These were also subjected to a second check. In fact, every step from checking to probabilities has been at least doubly checked.

Methods of Analysis

Primary interest inheres, of course, in the feature of salience and in the analyses which determine the extent of it, as well as the interrelations between salience in one place and that in another. The description of the method of computing salience ratios was given in detail in the first paper on terminal salience (17), and the mathematics of the evaluation of the ratios was described by Dr. Greenwood in an article accompanying that first report (5). Therefore, only a brief outline of the steps involved need be given here:

First, the distribution of hit-frequency through the run was concentrated in five sub-totals, one for each of the five segments of five trials each. This gave the distribution of the run as it will be dealt with in this article.

Second, the same distribution for the twenty-five trials in the run was recombined in a different way, within the segment itself, by taking the first trial in each of the five segments, and totaling these; and then taking the second trial in the segment for each of the five segments and arriving at a second total; and so on through the five trials of the segment. Thus, five different sub-totals were obtained, giving the distribution of hits in the *segment*. By cutting horizontally across the record sheet and taking the total for each position in the columns, we get the distribution in the segment in the form of five sub-totals, one for each trial. The vertical check, then,

would give us the distribution in the run, and the horizontal, the distribution in the segment.

As a third step, the distribution of deviations is found simply by subtracting from each of the sub-totals the expectation for the number of trials represented. A half-point (.5) is dropped from each deviation, according to Yates correction. [See Greenwood's article (5).] The standard deviation is known for each ($\sqrt{npq}$), and the CR is obtained for each sub-total. The CR squared becomes the chi-square with one degree of freedom. When the chi-square is obtained for each position in the segment, and for the five sub-totals in the run, the salience ratios may be obtained from these values. This ratio is the sum of the two end chi-squares (the first and fifth positions) divided by the sum of the three middle chi-squares (second, third, and fourth positions). This is the same for the segment and the run. If there is no salience, this quotient would come out exactly two-thirds. Theoretically, the ratios may vary all the way from zero to infinity.

The salience ratio, or SR, for the run is the RSR, and that of the segment, the SSR. These may be evaluated in terms of the probability of the occurrence of as great or greater SR's, but the primary feature of interest in these studies has been the interrelations between SSR's and RSR's. The extent of these interrelations has been determined by covariation statistics, for which a CR and the corresponding probability may be found, as described in the Greenwood article referred to above (5). The covariation statistic is applied likewise to interrelations between SR's in the two Divisions—Child and Adult.

In the covariation treatment it is assumed that the SSR's and RSR's are either independent, or that any interdependence that may exist is negligible. The question as to whether or not there is dependence has been considered in previous reports on salience, and the conclusion stated that the two measures are effectively independent of each other. An empirical check reported by Pratt, Humphrey and Rhine in the March issue of the JOURNAL (13) confirmed this position, since a computation of SR's on the basis of deviations secured by a random reshuffling of the hits of the Subseries in Series I gave only chance covariations on all the levels reported in the original article on this Series.

Results

Critical Ratio Method

As evaluated by the CR method, the 1,000 runs of 25 trials each are insignificant. The total deviation of + 56 gives a CR of only .89. The same verdict is applicable also to the Divisions taken separately, as well as to the Sections, Immediate and Delayed, whether these are taken by Divisions or by the combination of Divisions. So far, then, as total deviations are concerned, the results may be ascribable to chance. Table 1 presents these figures.

Table 1

Results Evaluated by the CR Method

Section	Child Division			Adult Division			Both Divisions		
	Runs	Dev.	CR	Runs	Dev.	CR	Runs	Dev.	CR
Immediate	204	+10	.35	285	+44	1.30	489	+54	1.22
Delay	237	+45	1.46	274	−43	1.30	511	+ 2	.04
Total	441	+55	1.31	559	+ 1	.02	1,000	+56	.89

The Salience Ratio Method

In the first report on terminal salience, it was suggested that the SR might be a useful basis for determining extra-chance relations in ESP results, and such an application was made in that instance. In the second report, the covariation of SR's became the sole basis of evidence of extra-chance factors in the results of the transoceanic DT tests reported. The same is true of the third article, in which covariation of SR's provides extra-chance evidence of precognition. The present report is a fourth instance of this application of covariation in the establishment of ESP relations, and again the SR becomes the sole basis of evidence against the chance hypothesis.

In Series IV, as was true more or less in Series I, II, and III, the SR's are not themselves statistically significant, with the exception of a few outstanding cases (notably Series I); but the relations to each other, as measured by the covariation statistic, are in several ways of grouping reliably different from expectation on the chance hypothesis. But while primary interest attaches to the interrelations between SR's, it will be of some advantage to present first in detail

the SSR's and RSR's obtained for the various subdivisions of the research reported in Series IV. These appear in Table 2. Therein will be found the values given separately for the Adult and Child Divisions, and in Parts A, B, and C the results, first by Subseries, then by Sections, and finally by Division totals.

For the benefit of those interested in further detail, an Appendix Table will be given, as has been done in each case of the earlier reports, providing the actual distribution of chi-squares from which each SR has been computed. We will furnish other information by correspondence, if it is desired.

Table 2

SALIENCE RATIOS AND THEIR PROBABILITIES

A. BY SUBSERIES

Subseries No.	Section	Adult Division d.f.	SSR	2P	RSR	2P	Subseries No.	Child Division d.f.	SSR	2P	RSR	2P
1	Immediate...	2,3	.07	.16	.90	.76	1	2,3	5.96	.12	4.28	.16
2	" ..	2,3	1.01	.70	.93	.75	2	2,3	.62	.97	.54	.96
3	" ..	2,3	1.23	.60	.43	.83						
4	" ..	2,3	.76	.86	2.08	.37						
1	Delay ..	2,3	3.36	.22	.13	.33	1	2,3	.17	.41	2.33	.33
2	" ..	2,3	.07	.16	.14	.37	2	2,3	24.12	.02	7.73	.08
3	" ..	2,3	.85	.80	.29	.64						
4	" ..	2,3	.27	.60	.05	.14						

B. BY SECTIONS

Immediate....	8,12	.45	.59	.78	.80		4,6	1.81	.26	3.32	.08
Delay.....	8,12	.58	.87	.22	.12		4,6	3.94	.07	3.29	.08

C. BY DIVISIONS

	16,24	.52	.62	.47	.48		8,12	2.54	.04	3.31	.013

The P-values have been doubled in all instances to allow for the evaluation of deviations in either direction from the approximate median of .67, which is the no-salience point; that is, either terminal salience or its opposite, middle salience, are of equal psychological interest, provided they are statistically significant.

Interrelations among SR's

In the earlier salience studies, there have been three types of interrelations mentioned: first, and of primary significance, the SSR-RSR relation, or that between the SR's of the segment and those of the run; second, the relation between the SR's in one Division, and the corresponding SR's in another—this has mainly consisted of interrelations between Child SR's and Adult SR's, or Child-Adult covariation; third, in the preceding report, interrelations were examined between SR's obtained from the deviation distributions of the Subseries pooled into larger groupings, such as the Series and Divisions, before obtaining the chi-squares. These SR's are not wholly independent of those obtained by the usual method of combining chi-squares, though they have considerable confirmatory value.

SSR-RSR Covariation

When the SSR's are paired with RSR's in the covariation treatment of Series IV, the result is definitely consistent with the findings of the previous work. First, as found before at the level of the *Subseries* (except in the transoceanic DT work), an insignificant CR of 1.70 is obtained; on the level of the *Sections*, however, the CR is 3.03, with a probability of .003; while on the level of the *Divisions*, the CR is 4.09, which gives a P-value of .001. These results can be found in Table 3. The corresponding SR's from which the computation has been made are in Table 2.

Table 3

COVARIATION BETWEEN SSR'S AND RSR'S

Subdivisions	Number of Pairs	CR	P
A. By Divisions	2	+4.09	.0010
B. By Sections	4	+3.03	.0031
C. By Subseries	12	+1.70	.044

In Series IV, as in the earlier Series, the covariation CR's for the SSR-RSR relations are positive. The three CR's of this investigation make up a total of eleven positive covariation CR's, with no negative ones. Further comparisons with the SSR-RSR covariations of earlier Series will be discussed below, where a complete summary will be offered.

Child-Adult SR-Covariation

In the establishment of extra-chance relations in the work showing salience, the Child-Adult covariation has been nearly as marked as SSR-RSR covariation. In Series I, II, and III, there were significant covariation CR's obtained from the Child-Adult SR's, though they were not quite as large as those which the SSR-RSR relation gave. A similar situation, one supporting the case against chance, is furnished in Series IV; namely, the covariation between Child SR's and Adult SR's is significant ($2P = .0018$) in one of the three orders of subdivisions. In this case, it is the Subseries which afford the significant CR, whereas only the larger subdivisions in the other Series gave significant CR's. This unusual feature will be taken up further in the discussion to follow. Table 4 gives the results of the Child-Adult covariation.

Table 4

COVARIATION BETWEEN CHILD SR's AND ADULT SR's

Subdivisions	Number of Pairs	CR	2P
A. By Divisions	2	−2.15	.038
B. By Sections	4	−1.72	.083
C. By Subseries	8	−3.25	.0018

It will be seen that this table offers a column headed "2P" instead of "P." This is due to the fact that the covariation CR's are negative, as they were likewise in Series II and III, though not in Series I. Since there is no *a priori* reason for excluding negative CR's, and since they may be fully as meaningful psychologically as positive ones, it is important to take note of them. In order to evaluate them properly, however, the P-values are multiplied by two.

SR's from Pooled Deviation

In the preceding paper on salience, the SR's were computed also from the pooled deviations; that is, with the deviation distributions of the Subseries pooled for each Division, before the chi-squares and salience ratios were determined. This method of measuring salience confirms the SSR-RSR relation found when the SR's were computed in the usual way. While not an independent method, it adds support to the primary findings. In Series IV the CR obtained from the

covariation between SSR's and RSR's obtained from the pooled deviations of the Divisions is + 1.87. This figure is not significant, having a P-value of .031; however, it is in line with the earlier work and supports the general findings, inasfar as it has weight.

Discussion

Alternative Hypotheses

Chance. The case against chance primarily rests upon the P-values of the SSR-RSR covariation. These values for the Divisions and the Sections are .001 and .003 respectively, with *a priori* grounds for the expectation of positive covariation. Confirmatory as this is of the work in Series II and III, as well as that of Series I, it has therein additional significance, just as it lends additional significance to the earlier findings themselves. The additional support against the chance hypothesis reported by the Child-Adult covariations, which on the Subseries level give a P-value of .00088 (which, corrected for negative sign, is .00176) is very substantial indeed. On the grounds of either of these two interrelationships, a rejection of the chance hypothesis would appear to be necessary.

Multiple-Calling Hypothesis. This hypothesis supposes that when a number of calls or predictions are checked against the same run of symbols or decks of cards, especially if the calls are made by the same subject, there may be call-patterns that will coincide with accidental distribution of symbols, so as to produce either unusually high or unusually low scores or score totals. Since, however, we are not concerned in this report with scores or score totals, this hypothesis would not here be applicable as ordinarily considered. It might be supposed, however, that in a similar way peculiar hit-distributions within the run might be effected. This multiple-calling hypothesis was discussed at some length in the preceding paper on salience, under the heading *"Call-Patterning and Preferences,"* in the course of which various arguments against its applicability were offered, as well as some pertinent experimental evidence.

In connection with Series IV, however, it can be shown quite as distinctly as it was in Series II and III, or Series I, that this hypothesis does not apply; for if it did, we should expect that the subdivision showing the largest number of runs would show the SR's with smallest P-values; those with the smallest number of runs

would give SR's with the largest P-values—that is, the least significant. This is definitely not the case. The smallest P-value by any Subseries is that given by Subseries No. 2 of the Delayed Section of the Child Division. The SSR gives a P-value of .02. The length of the Subseries is 127 runs. By combining the diagram of Figure 1 with the results given in Table 2, it may easily be seen that the second Subseries of the Adult Division, both in the Immediate and the Delayed Sections, are both larger than the Subseries of the Child Division just mentioned; yet both of these Adult Subseries have larger P-values instead of smaller. As a matter of fact, the Adult Subseries mentioned for the Immediate Section give P-values of .70 and .75. In a word, the hypothesis does not fit.

Numerous other conflicts can be found by combining the diagram and the table mentioned. To cite another instance, a comparison of the first two Subseries of the Adult Division might be given. Here we find a short Subseries of 40 runs, giving a P-value of .16 for the SSR, and a long Series of 150 runs giving, for the corresponding SSR, a P-value of .70—quite the opposite of what would be expected from the hypothesis of multiple-calling effect. A comparison of the P-values of the two Subseries of the Child Division shows the same thing again, whether it be taken from the Immediate Section or the Delayed. This hypothesis must therefore be rejected.

Other Counter-Hypotheses. With the precautions that have become a routine procedure in the conduct of experiments in this Laboratory, it hardly appears necessary to repeat discussion of the various counter-hypotheses of experimental errors—recording, checking, or computational errors—that clutter the controversial history of this field of investigation. If there are individuals in whose minds such questions arise, they may consult the authors as to the special safeguards concerning a particular alternative hypothesis. Furthermore, the previous reports on salience may be consulted for additional information on routine precautions.

The random superimposed cuts based upon the daily temperature figures of the weather reports offer assurance against any peculiar shuffling effects that might be imagined to have occurred. Likewise, this safeguarding cut presumably would help to exclude any unknown or unrecognized type of personal influence which the experi-

menter might conceivably exercise on the shuffling machine or the selection of the cards.

We are led, accordingly, to think these extra-chance results obtained are not explainable except by the hypothesis the experiments were designed to test: namely, that of extra-sensory perception expressing itself in peculiar hit-patterns that produce salience.

Salience

The U-Curve of the Hit-Distribution in the Run. The extra-chance character of the results depends solely on the SR method of measurement, and mainly upon the interrelations between the SR's rather than upon any individual SR. But since the investigation of salience grew out of the discovery of U-shaped curves in DT tests, followed by similar curves in PDT results, and since the preceding papers covering Series I, II, and III reported the finding of roughly U-shaped curves in both studies, it is of more than statistical interest to observe the shape of the curve given by the hit-distribution in the run for Series IV. This curve is shown in Figure 2. The solid line represents the distribution of pooled deviations for the run for the total data of Series IV. It is definitely similar to the previously observed U-curves.

The distribution of hits in the segment is a matter without so much precedent. What to expect in this matter is somewhat more doubtful. The curve shown by the broken line in Figure 2 is obviously not U-shaped. The corresponding curve for Series I and that for Series II and III have not been nearly so good approximations to U-shape as the distribution for the run. Nevertheless, they were much better than for Series IV.

We recognize that the U-curve is not in itself a significant finding, but it is nevertheless of considerable assurance to find it persisting throughout these Series. It would appear to indicate a generally predominant trend in respect to terminal salience running through the whole—a trend, moreover, that, as has been pointed out in earlier salience papers, bears a marked resemblance to curves of success in repeating memorized lists of words.

In a field of investigation in which it has been almost impossible to repeat a particular finding, the discovery that Series IV confirms the preceding Series in SSR-RSR covariation is of more than ordinary

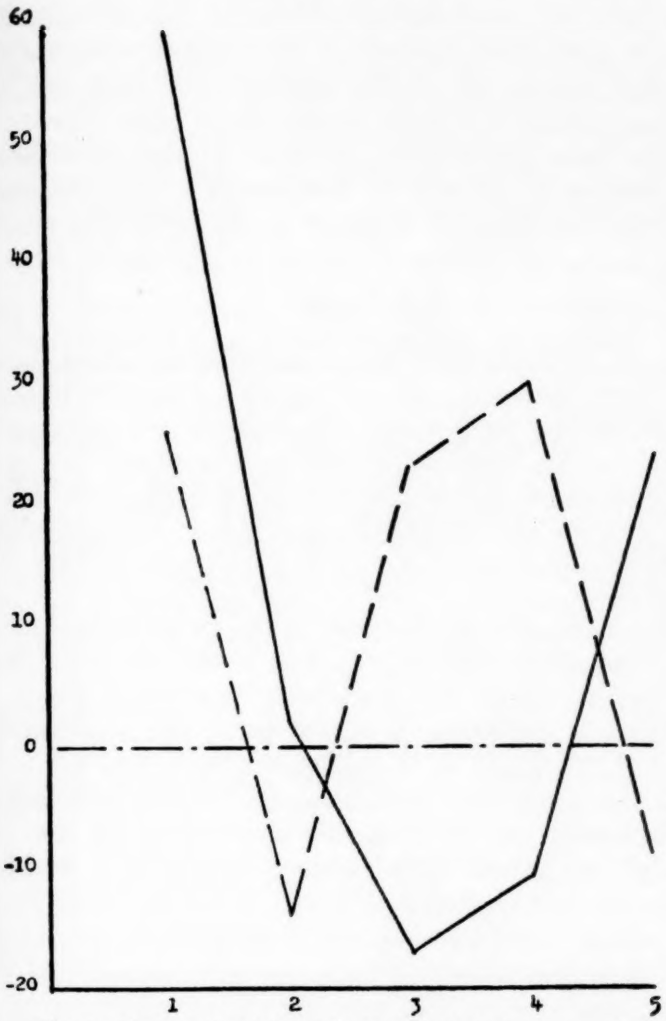

Fig. 2.—Curves of total distribution of deviation of run (solid line) and of segment (broken line) in Series IV. 1,000 runs; deviation = + 56.

significance. Moreover, the general pattern of covariation CR's for the different levels of subdivisions is equally impressive, as well as the continued positive relation between the RSR and the SSR. This whole SSR-RSR pattern for the four Series warrants summarizing, and we have assembled in Table 5 the covariation CR's and P-values

of the preceding reports on Series I, II, and III, and repeat those of Series IV given above, as well. It may be seen, then, that there is persistent insignificance of the Subseries covariation, and general significance of the covariation on the upper levels of subdivision—Division, Series, and Sections. In Series IV, there is no subdivision comparable to the Series in the earlier report. As a matter of fact, the whole subdivision of Series IV is somewhat different, but, as nearly as may be, is placed in comparable positions in Table 5.

Table 5

SUMMARY OF SSR-RSR COVARIATION FOR ALL FOUR SERIES

Subdivisions	Series I			Series II and III			Series IV		
	Pairs	CR	P	Pairs	CR	P	Pairs	CR	P
A. Divisions.....	2	+2.16	.019	2	+3.05	.0047	2	+4.09	.0010
B. Series.....	4	+3.37	.0019	4	+2.91	.0044	..	.....	.....
C. Sections.....	8	+3.21	.0020	8	+2.57	.014	4	+3.03	.0031
D. Subseries.....	28	+1.19	.12	28	+0.11	....	12	+1.70	.045

The reader who is interested in the peculiar problem created by the insignificance of the SSR-RSR covariation at the Subseries level may be referred to the preceding paper on salience, in which an attempt at explanation is offered under the paragraph heading "*SSR-RSR Relations*" in the Section on "*Discussion.*"

Covariation of Child and Adult SR's. In the preceding paper, the Child-Adult relation of the SR's was a negative one, and the suggestion was offered that the falling-off of the Adult Division SR's which led to this negative relation may have been due to the lack of adult interest in the test—perhaps to boredom; for the adults continued to produce low SR's, while the ratios of the Child Division went up considerably. It seems worth while to summarize the SSR's and RSR's for the four Series under both Divisions, and this is done in Table 6 below.

We have observed, as stated above under the heading of "*General Conditions and Attitudes,*" a better maintenance of interest in the test on the part of children than of the adult subjects, and there were mentioned some evidences of lack of enthusiasm on the part of the latter. The connection between these attitudes and the decline of SR's can, of course, be only suggested, not established. In this

Table 6

SALIENCE RATIOS OF THE FOUR SERIES
COMPARISON OF CHILD SR'S AND ADULT SR'S IN THE FOUR SERIES

Series	Child Division		Adult Division	
	SSR	RSR	SSR	RSR
I	.80	.84	1.24	.85
II	1.21	1.34	.75	.56
III	1.07	1.32	.28	.13
IV	2.54	3.31	.52	.47

connection, it may be of interest to compare the Child-Adult covariation between SR's, as it has run through the entire four Series. Accordingly, we have summarized in Table 7 the essential CR's and P-values to permit this comparison. When Series I appeared, there was at that time no evidence of a negative relation between the Child and Adult Divisions, and accordingly the P-values were not multiplied by two to correct for the use of both signs. Now that this has been necessary for the later Series, however, any comparison of these involving Series I should likewise, we think, require the correction on this Series as well. Accordingly, all P-values in the table have been doubled.

Table 7

SUMMARY OE CHILD-ADULT SR-COVARIATION FOR ALL FOUR SIGNS

Subdivision	Series I			Series II and III			Series IV		
	Pairs	CR	2P	Pairs	CR	2P	Pairs	CR	2P
A. Total	2	+1.60	.098	2	−3.27	.0061	2	−2.15	.038
B. Series	4	+3.35	.0050	4	−2.19	.033	..	.....	....
C. Sections	8	+2.61	.012	8	−1.50	.13	4	−1.72	.083
D. Subseries	24	+1.62	.10	22	−0.11	...	4	−3.25	.0018

By inspection of Table 6, it may be seen that the transition from the positive to the negative correlation between the Child and the Adult Divisions occurred in Series II and that the gap has become increasingly wider in the succeeding Series.

In Table 7, it may be noted that there is an exception to what has appeared to be a general rule in the fact that the Subseries covariation in Series IV is significant. We would suggest as a possible

explanation the fact that these are extraordinarily large Subseries, as may be seen by turning to the diagram in Figure 1. Only the first two Series of each Section could be used for the covariation study, since the Child Division has only two Subseries per Section. We have noted above in SSR-RSR covariation treatment that the larger the subdivision, in general, the larger the covariation CR. The same has been true, roughly, in the Adult-Child covariation in the first three Series. Now, since in Series IV we have extraordinarily large Subseries, we might, on the same logic, expect to get larger covariation CR's than we would if the Subseries were smaller.[2] It is, to our judgment, of no great immediate consequence as to whether this explanation is correct or not. It represents another side-problem, one which may well be solved in the course of the analyses under way which cover the material of these four Series.

Rather, the important point at this stage is the continued significance of the Adult-Child SR-covariation and the consequent support which this relationship offers to the extra-chance character of the results evidenced by the SSR-RSR covariation.

Psychologically, we are hardly ready as yet with the information we possess to offer more than conjectural concepts for the differences revealed by the negative covariation. It would appear, however, that we have here a combination of a psychological and a statistical method, which is perhaps more sensitive as an instrument of comparison than have been the more customary statistical methods, at least for the particular situations with which we are dealing.

Precognition Hypothesis

Inasfar as the results are extra-chance in character, which is to say inasfar as they are significant, the effect of salience can only be accounted for under the conditions of the experiment by supposing that some knowledge of the target cards must have been possessed by the subject, however unaware of it he may have been and however slight and dispersed through the Series was its effect. In other words, some of the subjects must have known, or acted as if they

[2] This has nothing to do with the Multiple-Calling Hypothesis discussed above under "*Alternative Hypotheses.*" This relation of size of subdivisions to covariation-CR is clearly unrelated to the question of the number of runs checked against a given deck of cards, which was the question under discussion in connection with that hypothesis.

knew, something about the order of decks of cards before such order was yet established. The fact that the order of the decks was determined by processes and conditions designed to be thoroughly random classifies the knowledge referred to as precognition. Therefore, unless there is some fundamental error in the experiment or in the interpretation of the results, it would appear to follow that significant knowledge of uninferable random future events has been shown to occur.

However inconspicuous the extent of precognitive knowledge shown, and however obscure the ultimate explanation may be, we (or our readers) must either find an essential weakness in the research or else be forced to recognize the obviously far-reaching consequences of the conclusion to which they would point—that cognition may transcend time, as it has been shown to do with space (19).

Length of Time in Precognition. Concluding as we do that, to our present knowledge, there is no alternative to the precognition hypothesis as an explanation of the results reported above, it is now in order to examine the results of the Immediate and the Delayed check with a view to comparing the effects of a short time with those of a longer period upon performance in precognition tests. Most of the tests in the Immediate Check Section had a period of one or two days between the filling-out of the record sheet and the checking of it. We have, then, effectively, a two-day versus a ten-day comparison, much as in the previous report. The difference here is that the subjects did not know which of their runs were two-day and which were ten-day.

Nevertheless, the fact that the Immediate Section gives a covariation-CR for the SSR-RSR relation of 1.49, while that for the Delayed Section is 2.82, shows that again, as for the preceding study, the longer time is no limitation upon performance. In other words, if precognition occurs, it occurs quite as well over a ten-day period as over a two-day. The covariation-CR for the Delayed Section is independently significant, having a P-value of .0058.

But that which perhaps offers the greatest enlightenment in the present phase of our investigation is the fact that the subject need not know or even give his attention to the question of when the check-up is to be made. He succeeds (as measured by SR-covariation) not only when he knows when the checking of a given run will be

done, but quite as well when he does not. This would suggest that the subject, in filling out the record sheet, somehow short-circuits the many pathways which we might rationally suppose would be involved in the prediction of the order of symbols. It would seem that he somehow is able to put down simply what is going to be on the record sheets without following the intermediating train of events. This hypothesis may, for the time being, be regarded as at least favored somewhat by the results, and further study may reveal a way to put the matter to a much more crucial test than this has done. It is perhaps something of an extension of Foster's Diametric Hypothesis.[3]

Comparison of Child and Adult Divisions as to Evidence of Precognition. It would appear possible to use the SSR-RSR covariation as a basis for comparison of the evidence of precognition given by the two Divisions. From this evidence it would appear that the case for precognition rests almost entirely, so far as Series IV is concerned, on the Child Division; for, as Table 8 shows, this Division gives significant covariation-CR's for both the Section and Subseries levels. Indeed, the probability for the Section level is the smallest which has been given in any of the covariation studies of salience to date. It is in marked contrast with the figures for the Adult Division. The latter are quite insignificant.

Table 8

COMPARISON OF CHILD AND ADULT DIVISIONS AS TO SSR-RSR COVARIATION

Subdivision	Adult Division			Child Division		
	Pairs	CR	P	Pairs	CR	P
A. Sections.....	2	+.12	...*	2	+4.18	.00058
B. Subseries....	8	−.10	...*	4	+2.80	.0052

*Negligible.

It was pointed out above that taking the two Divisions together, the Delayed Section contributed a much larger, indeed a significant, covariation-CR in the SSR-RSR relation. However, taking the Child Division alone, there is no sharp distinction between the Immediate and the Delayed, although the Delayed still gives a somewhat larger covariation-CR. The SSR-RSR covariation-CR's for the Immediate

[3] See this JOURNAL, vol. 4, 1940, pp. 325-328.

and Delayed Sections of the Child Division respectively are 1.93 and 2.01.

Summary

Once more an ESP experiment has been conducted under conditions which would not appear to permit error in favor of the hypothesis, and the results are found to be extra-chance in character. The significant probabilities obtained are derived from the salience ratio method, specifically, from certain interrelations between salience ratios evaluated by the covariation statistic. The Multiple-Calling Hypothesis is considered and shown to be inapplicable. The results are regarded as evidence of extra-sensory perception.

At the same time the experiment has been one that involved precognition, again with mechanical selection of the card order against which the predictions were made; again, as in the case of the preceding report, a further safeguarding cut was introduced which was determined by the figures from maximal and minimal temperature readings. With the methods of randomization used, no recognizable alternative is available to explain the results, except the precognition hypotheses itself.

Further, this has been another experiment in precognition conducted with a view to testing whether length of time between prediction and verification has a determining influence upon the results. As in the preceding study, the ten-day period was set off against a shorter period of one or two days. In this, as in the earlier instance, the better results were obtained under the conditions of the longer period. They are not significantly better, and do not warrant the view that the longer time is necessarily an advantage, but they do justify the conclusion that the longer time is not a limiting condition. And, most suggestive feature of all, the effect has been obtained without the subject knowing whether the run he was participating in would be checked after two or after ten days time. The determination of that, too, was left to the future.

The most consistent and reliable support for the extra-chance character of the results in this study, as in the preceding ones of comparable character, lies in the significant covariation found between SSR's and RSR's; that is, salience ratios of the segment and the run. The relationship has been throughout positive, thus far, and has been

significant in the larger groupings or subdivisions of the data, and insignificant on the Subseries level except in the transoceanic DT Series. This represents an unusual pattern of continuity against the erratic character of the results as measured by the CR's of the total deviations.

But the extra-chance nature of the results is confirmed by the hardly less significant probabilities deriving from the covariation of Child SR's and Adult SR's. This, too, represents an independent measure of extra-chance SR interrelations, and the fact that this line of evidence is also confirmatory of similar findings in preceding studies gives it still further weight.

Another U-curve for the hit-distribution in the run is added to the long series now on record from DT and PDT studies. The corresponding curve for the segment, however, is not U-shaped.

The Child SR's have continued to rise through the Series reported in this paper, while the Adult SR's remained at a low level. The gap between the two is, therefore, wider than before. Accordingly, the Child-Adult relationship between SR's is a negative one.

Furthermore, it is clear that the Child Division is mainly responsible for the extra-chance nature of the SSR-RSR relationship. Probabilities for this relation in this Division reach the low mark of .0004. The corresponding probabilities for the Adult Division are insignificant.

The Delayed or Ten-Day Section is independently significant in its SSR-RSR covariation, and this, along with the fact that the Child Division likewise stands alone in statistical significance, permits some localization of the functioning of the extra-chance factor responsible. However, at this stage of the investigation, it is evident that whatever importance the work reported may have lies not so much in these differences as in the gross findings mentioned in the first three paragraphs of the Summary. In a word, then, ESP occurs, and, up to a certain point, at least, is independent of time as ordinarily conceived.

APPENDIX

Table 9

THE CHI-SQUARES, SALIENCE RATIOS (AND THEIR PROBABILITIES) FOR SERIES IV

A. IN THE RUN

CHILD DIVISION

Section	Sub-series	Total Runs	Total Dev.	Chi-Squares per Segment 1	2	3	4	5	RSR	d.f.	2P
Immediate	1	79	+13	6.66	.....	1.44	.19	.32	4.28	2.3	.16
"	2	125	− 3	.30	.....	.....	.56	.....	.54	2.3	.96
Immediate Total		204	+10	6.96	.....	1.44	.75	.32	3.32	4,6	.08
Delayed	1	110	+15	1.02	.23	.14	.83	1.77	2.33	2,3	.33
"	2	127	+30	.71	.06	.20	.....	1.30	7.73	4,6	.08
Delayed Total		237	+45	1.73	.29	.34	.83	3.07	3.29	4,6	.08
Child Division Total		441	+55	8.69	.29	1.78	1.58	3.39	3.31	8,12	.01

ADULT DIVISION

Immediate	1	40	+23	1.77	.94	.64	.38	.....	.90	2,3	.76
"	2	150	+ 1	.10	1.74	1.51	.25	3.17	.93	2,3	.75
"	3	35	+17	.01	3.24	.22	.08	1.51	.43	2,3	.83
"	4	60	+ 3	.13	.26	.05	.05	.62	2.08	2,3	.37
Immediate Total		285	+44	2.01	6.18	2.42	.76	5.30	.78	8,12	.80
Delayed	1	45	−23	.....	3.06	1.17	.....	.56	.13	2,3	.33
"	2	154	+ 1	.....	.02	.25	.10	.05	.14	2,3	.38
"	3	35	−27	.08	.22	.44	6.50	2.02	.29	2,3	.64
"	4	40	+ 6	.01	.01	.38	.01	.01	.05	2,3	.14
Delayed Total		274	−43	.09	3.31	2.24	6.61	2.64	.22	8,12	.12
Adult Division Total		559	+ 1	2.10	9.49	4.66	7.37	7.94	.47	16,24	.48

B. In the Segment

Child Division

Section	Sub-series	Total Runs	Total Dev.	Chi-Squares per Trial 1	2	3	4	5	SSR	d.f.	2P
Immediate	1	79	+13	6.66	.48	.19	.48	.19	5.96	2,3	.12
"	2	125	− 3	.42	1.82	2.10	.12	2.10	.62	2,3	.97
Immediate Total		204	+10	7.08	2.30	2.29	.60	2.29	1.81	4,6	.26
Delayed	1	110	+15	.35	.....	.23	2.07	.03	.17	2,3	.41
"	2	127	+30	10.37	.41	.02	.....	.....	24.12	2,3	.02
Delayed Total		237	+45	10.72	.41	.25	2.07	.03	3.94	4,6	.07
Child Division Total		441	+55	17.80	2.71	2.54	2.67	2.32	2.54	8,12	.04

Adult Division

Immediate	1	40	+23	.19	1.32	2.82	.94	.19	.07	2,3	.16
"	2	150	+ 1	.35	.....	.92	.92	1.51	1.01	2,3	.70
"	3	35	+17	.44	.22	.72	.....	.72	1.23	2,3	.60
"	4	60	+ 3	.....	.42	.13	.....	.42	.76	2,3	.86
Immediate Total		285	+44	.98	1.96	4.59	1.86	2.84	.45	8,12	.59
Delayed	1	45	−23	2.50	.....	.06	.85	.56	3.36	2,3	.22
"	2	154	+ 1	.17	.25	4.49	.35	.17	.07	2,3	.16
"	3	35	−27	1.08	.01	2.59	.44	1.51	.85	2,3	.80
"	4	40	+ 6	.07	.64	1.32	.64	.64	.27	2,3	.60
Delayed Total		274	−43	3.82	.90	8.46	2.28	2.88	.58	8,12	.87
Adult Division Total		559	+ 1	4.80	2.86	13.05	4.14	5.72	.52	16,24	.62

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. Analysis of a large chance control series of ESP data, *J. Parapsychol.*, 1938, **2**, 138-146.
5. ———. The statistics of salience ratios, *J. Parapsychol.*, 1941, **5**, 245-249.

6. HUMPHREY, BETTY M. AND PRATT, J. G. A comparison of five ESP test procedures, *J. Parapsychol.*, 1941, **5**, 267-292.
7. HUTCHINSON, L. Variations of time intervals in pre-shuffle card-calling tests, *J. Parapsychol.*, 1940, **4**, 249-270.
8. MACFARLAND, JAMES D. Discrimination shown between experimenters by subjects, *J. Parapsychol.*, 1938, **2**, 160-170.
9. MARTIN, D. R. AND STRIBIC, F. P. Studies in extra-sensory perception: I. An analysis of 25,000 trials, *J. Parapsychol.*, 1938, **2**, 23-30.
10. ———. Studies in extra-sensory perception: II. An analysis of a second series of 25,000 trials, *J. Parapsychol.*, 1938, **2**, 287-295.
11. ———. Studies in extra-sensory perception: III. A review of all University of Colorado experiments, *J. Parapsychol.*, 1940, **4**, 159-248.
12. PEGRAM, MARGARET H. Some psychological relations of extra-sensory perception, *J. Parapsychol.*, 1937, **1**, 191-205.
13. PRATT, J. G., HUMPHREY, BETTY M. AND RHINE, J. B. A check on salience relations in ESP data, *J. Parapsychol.*, 1942, **6**, 44-51.
14. RHINE, J. B. Experiments bearing on the precognition hypothesis, *J. Parapsychol.*, 1938, **2**, 38-54.
15. ———. Experiments bearing upon the precognition hypothesis. III. Mechanically selected cards, *J. Parapsychol.*, 1941, **5**, 1-57.
16. ———. Extra-Sensory Perception. Boston: Bruce Humphries, 1934.
17. ———. Terminal salience in ESP performance, *J. Parapsychol.*, 1941, **5**, 183-244.
18. ———. Evidence of precognition in the covariation of salience ratios, *J. Parapsychol.*, 1942, **6**, 111-143.
19. RHINE, J. B. AND HUMPHREY, B. M. A transoceanic ESP experiment, *J. Parapsychol.*, 1942, **6**, 52-74.
20. RHINE, J. B., SMITH, B. M. AND WOODRUFF, J. L. Experiments bearing on the precognition hypothesis: II. The role of ESP in the shuffling of cards, *J. Parapsychol.*, 1938, **2**, 119-131.
21. RHINE, J. B., PRATT, J. G., STUART, C. E. AND SMITH, B. M. WITH GREENWOOD, J. A. Extra-Sensory Perception After Sixty Years. New York: Henry Holt and Co., 1940.
22. SALTMARSH, H. F. Report on cases of apparent precognition, *Proc. S.P.R.*, 1934, **42**, 49-103.
23. STUART, C. E. An analysis to determine a test predictive of extra-chance scoring in card-calling tests, *J. Parapsychol.*, 1941, **5**, 99-137.
24. TYRRELL, G. N. M. Further research in extra-sensory perception, *Proc. S.P.R.*, 1936, **44**, 99-168.

A SECOND STUDY OF THE EFFECT OF TEMPO RATES OF MATCHING

By CHARLES E. STUART, *Duke University* and
BURKE M. SMITH, *U. S. Navy*

ABSTRACT: In a previous experiment, it had been found that subjects matching cards at a preferred rate of movement made higher ESP scores than when they matched the cards at a non-preferred rate. In the present investigation, a similar comparison was carried out, with the exception that the rate was emphasized by a flashing light and 50 trial runs were used instead of 25.

The experimental plan was broken by an unforeseen interruption which separated the results into two series, carried out a year apart, and introduced uncontrolled motivational differences.

It was found that, in the first series, the results verified the initial predictions that extra-chance scoring would occur at the preferred rate and that chance scoring would occur at the non-preferred rates. In the second series, the scoring was predominantly negative (i.e., below mean chance expectation) and showed no consistent favoring of rate. The two series were significantly different and thus the results supported the hypothesis that marked motivational differences accounted for the failure of the second series to display a tempo effect.

In 1938, a study of the effect of rate of matching upon ESP scores was presented in this JOURNAL.[1] In a series of experiments with matching procedures, it was found that subjects scored significantly when their rate of movement was fixed to correspond to their normal "tempo," a tapping rate timed at the beginning of the experimental session. The same subjects scored at a chance level

[1] C. E. STUART. The effect of rate of movement in card-matching tests of ESP. *J. Parapsychol.*, 1938, **2**, 171-183.

when their rate of movement was markedly faster or slower than the tempo rate. It was suggested that the comfortable rate of movement at the tempo rate enhanced a feeling of successful effort and permitted predominant attention to the ESP task of correctly matching the cards. On the other hand, the non-tempo rates required attention to the rate itself in the matching task and thus distracted primary attention from the ESP matching.

This experiment seemed to provide a neat demonstration of differential scoring conditions, but much of its force was lost because the experimental procedure did not enforce certain objective precautions of recording which subsequently became the topic of extensive controversy. Thus an interesting psychological correlate of the ESP process became well-nigh lost among the needs of demonstrating ESP under more and more rigorous and limiting conditions.

But fortunately, the tempo effect presented a definite experimental advantage. It was easy to control and, indeed, seemed exceptionally well designed to permit manipulation in a rigorously planned experiment. So we set out to plan a repetition study of the tempo effect under well-tested precautionary conditions. Our initial interest was in finding a readily repeatable demonstration of differential ESP scoring. The tempo effect seemed to be the kind of objectively controllable condition that would be ideal for such a demonstration. After some preliminary trial-and-error sessions, a rigorous procedure was adopted.

Procedure

When the subject came into the experiment room, he was seated at one side of the experimental table. After due social orientation, he was instructed to tap with a pointer on the table, the tapping to be regular and in three-quarter time. If he could not attain the rhythm, it was demonstrated to him, but irregularly. After he had attained a regular beat, Stuart timed the beat for thirty seconds with a stop-watch. This rate was entered in the record as the subject's normal tempo.

Smith, seated on the opposite side of the experimental table, then explained what the subject was to do and, if the subject had never been in an ESP test situation before, demonstrated the card-matching procedure.

On the table was an STM screen. The five key cards underneath the screen were enclosed in black opaque envelopes. Just above the subject's eye level was a forty-watt red lamp in a reflector. The reflector was so arranged that the light might be turned directly in the subject's face or turned away from him on to an area of the screen covered with rough black paper. The red light was connected with a mercury-cup contact on a metronome placed on a table just behind the experimenter's side of the experimental table. (During the matching runs, Stuart was seated at the metronome table and was screened from the subject.) When the metronome was started, the light flashed rhythmically at the rate set.

The subject was instructed to tap the key cards to indicate where the experimenter (Smith) should lay each card in turn. He was told to tap a key card each time the light flashed and to make his designations with an easy rhythmic motion of his hand and arm.

Each test run consisted of the matching of a shuffled deck of 50 ESP cards with symbols equally distributed.

Each subject-session consisted of 16 such test runs. The first two runs were matched at a non-tempo rate. The second two runs were matched at the tempo rate. The same order of presentation was maintained for the remaining runs. The first run was done with the signal light flashing in the subject's face (i.e., with the light up); the second run with the light directed on the screen (i.e., with the light down). This strict alternation was maintained throughout the 16-run session.

The tempo runs were matched at a rate equal to the normal tapping rate estimated at the beginning of the experimental session. The non-tempo rate was calculated at either 3/2 the tempo rate for "fast non-tempo," or 2/3 the tempo rate for "slow non-tempo." (For a tempo rate of 60, the slow non-tempo rate would be 40; the fast non-tempo rate would be 90.) Either a fast non-tempo rate or a slow non-tempo rate was used through a given subject-session. Each subject, with one exception, took part in two experimental sessions, one in which the fast non-tempo rate was used, the other in which the slow non-tempo rate was used.

After each run and before any checking was done, the subject was asked to guess what his score would be after the checkup. (These

estimates are not used in this report.) Then the STM screen was removed from the table. Both Stuart and Smith laid out the matched cards face up in a row before the corresponding enclosed key cards. The enclosed cards were then taken one by one from the envelopes and the actual correspondences counted by both experimenters and the subject. The subject was motivated to count accurately by the promise of a twenty-five cent reward if, as each experimenter independently announced his total count, he could show that both had erred. (Actually no rewards had to be paid. Even a difference in count by the two experimenters was rare.) The experimenters recorded the run scores in separate record books.

Thirteen subjects took part in the experiment. Twelve of these served for two sessions and one for a single session, thus providing 25 subject-sessions in all. All the subjects were adults. Most were undergraduate or graduate students; the exceptions will be noted later. None of them were subjects who had shown any previous excellence in ESP tests, and for most it was their first card matching test. We endeavored to utilize only curiosity or general willingness to please the experimenters as the essential motivation. Our endeavors to make relation casual and emotionally unstrained were, however, not always successful.

Predictions and Planned Variables

In the previous observations by Stuart on the effect of tempo and non-tempo rates upon scoring success, the scoring average per 25 calls was about 5.5 at the tempo rates. Evidence was presented to indicate that this did not point to a facilitation of scoring by the tempo rate but that, rather, this was an average to be expected under the conditions, and non-tempo rates merely inhibited the maintenance of this normal ability.

Since the experimental set-up of the present experiment was less flexible than the previous work and since the matching task required response to enclosed cards rather than to the ESP symbols themselves, we predicted a lower scoring average for the present work. How much lower, it was impossible to estimate; but we assumed hopefully that it would not be below 5.3 per 25 trials. The rest of the experiment was designed to conform to this arbitrary expectation.

In particular, the nearest round figure in number of trials at a

5.3 scoring average that will give a clearly significant critical ratio is 10,000. Since it was desirable to have an equal number of non-tempo trials, the whole experiment must include 20,000 trials. With 50 card decks, this would require the matching of 400 decks. Since it is practicable to work about an hour for a given session, the experiment could be accomplished in 25 experimental sessions of 16 deck-matchings each.

Inasmuch as the differential results of the previous investigation had been attributable to the increased attention required in order to adjust to the non-tempo rates, we proposed to add further distraction to the situation, the flashing light. This was intended to have an annoyance value enough to make the "light down" condition more comfortable for the subject than the "light up" condition. We hoped that the annoyance effect would be great enough to produce negative scoring. (In practice, the subjects reported no annoyance at the light, but a few occasionally shaded their eyes from it.)

Our predictions in regard to the various conditions were: 1. Significant scoring at the tempo rates and chance scoring at the non-tempo rates. 2. Negative scoring at the non-tempo rates with the light up in the eyes of the subject. 3. Higher scoring at the tempo rate with the light down than with the light up.

Unplanned Variables

When the present experiment was planned, we made the reasonable assumption that, other things being equal, the results should show a typical difference in scoring at tempo and non-tempo rates. We found ultimately, however, that other things did not tend to be equal at all and that, indeed, uncontrolled variables tended finally to dominate the results.

We desired to keep motivation relatively constant by imposing no constraint or persuasion upon subjects in getting them to participate. In general, we felt that we succeeded to a marked degree. But the exceptions were, unfortunately, critical.

Our first experience with motivational difficulty was accidental. Miss K. had made an appointment for a session and found she could not come. With an admirable sense of responsibility, she sent a Miss L. as a substitute. Miss L. seemed a bit unfriendly and proceeded to make some of the lowest run scores we had observed. Her total

deviation for the session was —20. We were less concerned, however, with her scores as they bore upon our experiment than with the indication they gave of some strong negative attitude toward the situation. In her second session, she seemed much more socially expressive and scored a total deviation of +20. We were relieved to think that there was nothing intrinsically wrong with the test set-up. And as she was leaving, we proposed that her work had been so interesting that we would appreciate her returning for further experiment. She assured us, with extravagant firmness, that she would not come back because she had "made up her points." Questioning elicited the background story: She was a freshman pledge in the sorority of Miss K. Punishment for minor rule infractions was effected by giving the erring pledge "points" which could be worked out by various extra-academic services for her guiding sisters. Her participation in our experiment was just such a penance, exacted much against her will. It thus ceased to be a mystery why our well-intentioned expressions of encouragement and approval directed to her performance had fallen a little flat!

A second accidental feature of the experiment was that in the late spring of 1940, one of us (Smith) was called suddenly to a teaching post in a college at some distance away. We planned to get together in the summer to complete the work, but these plans never matured. Thus, our next opportunity for working together occurred a full year later, in the summer of 1941.

At this point, we had to consider another type of motivational variable; namely, the experimenters' own attitudes. For when we had ceased work in 1940, the experiment had verified our predictions to a significant degree. The tempo runs gave significantly greater than chance deviations, and the non-tempo runs were in the chance range. Our only interest in continuation was to round the work out in the orderly manner planned. Opposed to the latter interest was the fact that we both were primarily occupied with other research projects.

These issues were brought out in our approaching Dr. J. B. Rhine for counsel and discussion. His advice was that, since the 1940 work was completely sound in terms of any question of sampling, we should take these motivational variables very seriously. Our compromise decision was to complete the plan of the experiment, but

whatever the outcome, to make expressly clear in our report the existence of this choice-point, to present the issue involved, and to treat the results statistically with reference to the hypothetically changed conditions.

When we resumed experimentation in 1941, our first subject yielded markedly low scores. She appeared nervous and somewhat ill at ease in the experimental situation. In discussion of the possibility that we as experimenters might be responsible for the increased frustration symptoms, we decided finally to try to avoid the adverse social factors by using as subjects persons who were closely associated with us. These people would, we predicted, be able to make this adjustment wholly in terms of the test without the distraction of meeting a social situation as well. At the time, this seemed a wholly plausible supposition. In retrospect, however, it was evident that we were only exchanging one difficult motivational problem for another. That is, the later subjects evidenced a much keener sense of responsibility to the experimenters for their successes and failures.

One further minor variation should be noted. It was planned that the thirteenth subject have only one session, in which both fast and slow non-tempo rates should be used in order that all conditions might include an equal number of observations. It so happened that our ninth subject could not return for his second session. We decided to retain his record and give the thirteenth subject the customary two sessions. Since the ninth subject had no runs at a slow non-tempo rate, the totals are not quite equal for the number of runs under the various conditions.

Results

The total experiment yielded 400 50-trial runs, the equivalent of 800 25-card runs, in which the total deviation was +45, a deviation negligibly different from chance.[2]

The primary comparison was between runs at the tempo rate and runs at non-tempo rates. The total for the tempo rate was 400 runs with a deviation of +59. The total non-tempo rates gave 400 runs

[2] The customary ESP run consists of 25 calls; and since our evaluation in terms of the binomial hypothesis is not affected thereby, we will express the results in terms of 25-card runs to avoid confusion for those familiar with the literature. It need simply be remembered that one experimental run was the statistical equivalent of two "runs" as we use the term here.

with a deviation of —14. This difference is in the direction predicted but is not statistically significant.

The other major variation, runs with the flashing light up in the face of the subject compared to runs with the light turned down on the screen, gave similar inconclusive results. With the light up, 400 runs yielded a deviation of —6, with the light down, 400 runs gave a deviation of +51.

These deviations, apportioned to the 200 runs under each of four conditions, were: non-tempo, light up: — 46; non-tempo, light down: +32; tempo, light up: +40; tempo, light down: +19. The only suggestive difference here is that between tempo and non-tempo runs with the light up. The critical ratio of the difference is 2.15. A remarkable feature is that with the light down there was no difference between the tempo and non-tempo scores.

In Table 1 are presented the deviations by subjects for all the conditions. The first column designates the subject, the second column gives the tempo rate as measured at the beginning of the session. The remaining columns give the deviation per 8 runs for the up and down positions of the light with fast non-tempo, tempo, and slow non-tempo rates. The dotted line in the table marks the separation between the observations made in 1940 and those made in 1941.

A separate treatment of the 1940 results (which, it should be recalled, constitute a permissible statistical sample) yields the following deviations with 136 runs under each category: non-tempo, light up, —30; non-tempo, light down, +32; tempo, light up, +53; tempo, light down, +37. The total non-tempo deviation is +2 for 272 runs, the total tempo deviation is +90 for 272 runs, an average of 5.01 and 5.33 respectively. (This result is a very close approximation to our initial prediction.) The tempo deviation has the significant CR of 2.73. The difference, while suggestive, is not statistically significant. It is evident that the difference is due almost wholly to scoring in the "light up" condition. The difference between non-tempo, light up, and tempo, light up, gives a significant critical ratio of 2.52.

The whole experiment, evaluated by separate treatment of the 1940 and 1941 data, yields approximately similar results. A combi-

Table 1

TOTAL EXPERIMENT: DEVIATIONS BY SUBJECTS AND CONDITIONS

Subject	Tempo Rate	Deviation per 8 Runs					
		Fast		Tempo		Slow	
		u	d	u	d	u	d
GA	52	− 2	+ 1	+ 6	− 1		
	54			+ 8	+ 7	− 1	− 5
FA	58	− 9	− 9	− 8	+ 6		
	58			+ 4	+ 7	+ 3	+ 6
KA	60	− 1	+ 3	0	+ 7		
	56			+ 4	− 2	+ 2	+ 9
JO	48	− 6	0	+ 6	− 2		
	54			+ 4	− 8	− 3	+ 1
SE	68	− 7	+ 6	+ 4	+ 1		
	58			− 6	+ 5	0	+ 6
JR	68	+ 6	− 2	+ 6	+ 3		
	70			+ 3	+ 1	+ 3	+ 5
HU	54	+ 5	− 4	+ 7	− 3		
	64			− 5	+ 9	0	− 3
DE	66	− 5	+ 5	+ 8	+ 3		
	68			+ 9	+ 3	− 7	+ 5
BO	78	− 8	+ 8	+ 3	+ 1		
MY	60	− 2	− 6	− 9	− 4		
	60			+13	− 4	− 3	+ 3
WS	44	− 3	− 1	+ 1	+ 3		
	44			− 4	+ 7	+ 1	− 5
PH	48	0	+10	− 6	− 8		
	56			0	− 2	− 2	− 3
ST	45	− 3	+ 4	− 5	+ 2		
	44			− 3	−12	− 4	− 2
Total	..	−35	+15	+13	+ 8		
				+27	+11	−11	+17

nation of critical ratios gives a negligible chi-square for the non-tempo runs and a significant chi-square for the tempo runs. This evaluation is tabulated in Table 2.

The scoring average at the tempo rate was markedly different in the two periods. In 1940, it was 5.33 hits per run; in 1941, it

Table 2

EVALUATION OF RESULTS BY COMBINING 1940 AND 1941 SESSIONS TREATED SEPARATELY

	Runs	Dev.	CR	CR^2	P
Non-tempo					
1940.....	272	+ 2	.06	.00	
1941.....	128	−16	.71	.50	
Total.....	...	....	...	.50	.80
Tempo					
1940.....	272	+90	2.73	7.45	
1941.....	128	−31	1.37	1.88	
Total.....	...	....	....	9.33	.01

was 4.76 hits per run. The difference is significant with a CR of 2.68.

* * * * *

The differences between the actual outcome of this experiment and the prediction of the original plan are enlightening. From one standpoint, it might be argued that there is nothing to explain, since the experiment as planned gave no significant results. But a certainly more fruitful view is that the investigator in this field is not yet in the position to renounce any observation of which he has not fully planned the control.

The expectation was that significant scoring would take place at tempo rates and that chance scoring would result under the imposition of non-tempo rates of matching. It was considered probable that the attention required to retain rhythmic activity at a non-favored rate of movement would distract attention from the primary ESP task of correctly designating the matched card. The supposition was that a flashing light with annoyance value would enhance the latter distraction.

The 1940 series seemed to verify these expectations. Significant scoring did take place at the tempo rates, and chance scoring at the non-tempo rates. That the non-tempo, light up, condition gave a negative deviation would suggest that the flashing light did enhance the distraction of the non-tempo rate. But it is apparent that the light up condition at the tempo rate gave the highest deviation, a

fact which suggests that the flashing light might equally have enhanced the experience of efficient movement at the tempo rate. It would be interesting to investigate whether other means of accentuating the given rate would have a comparable effect.

The causes of the predominantly negative scoring in the 1941 series are probably as complex as the motivational factors suggested when we resumed the experiment at that time. It is notable that the widest deviations occurred during that series (see Table 1) with +10, −12, and +13 being recorded for 8-run groups. But we insist that the difference in scoring in the two series is not properly interpretable as an accidental variation within a totally chance series. The interruption was a fortuitous event not dependent upon the observed scores. So the probability that the two series actually constitute an accidentally divided population of scores is about .003, the probability interpretation of the difference between tempo scoring in the two series.

The fact that the uncontrolled motivational effects cut across the experimental variables emphasizes again the importance of those intangible factors in the experimental situation that are summarized under "experimenter-subject relations." As experimenters, we approached each session with the best of will, yet frequently could do nothing but watch frustrations develop in the subject's performance. We had subjects who obviously enjoyed the matching task, others who manifested anxiety. Some subjects "tried too hard," but how were we to effect their adjustment to an intuitive judgment of that sort while the metronome and the time of the experimental period ticked on inexorably? It is from this welter of performance problems that the insights leading to real control of ESP must come eventually.

Conclusion

In this repetition of earlier work in which subjects matched cards with greater success at their normal tempo rates than at markedly different non-tempo rates, it was found that scoring was significantly high at the tempo rate and negligibly different from chance at the non-tempo rates.

The utilization of a flashing light with supposed annoyance effect seemed to emphasize the difference between tempo and non-tempo

scores. It is suggested that the flashing light simply emphasized the effect of the rate of movement, rather than providing any definite effect in itself.

The significant effect of uncontrolled variables, supposedly of a motivational type, indicates that the matching rate is not alone an important enough factor to necessarily produce differential ESP scoring.

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.*