A REVIEW OF RECENT CRITICISMS OF *ESP* RESEARCH

CHARLES E. STUART

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

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

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

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

Criticism of Experimental Methods

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

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

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

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

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

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

This experimental result is ignored in the discussion and conclusion. The *unreliable* positive deviation is extrapolated to show that it *would be* significant in 10,000 runs, a significance due simply to the fact that the act of extrapolating assumes the original average to be reliable. The conclusions are stated that "these results are not striking but they do illustrate the possibilities of finding selected agents who may make these unnoticed errors consistently."[2] . . . "Whenever experimental procedures in telepathy and clairvoyance are so devised that unnoticed recording errors are permitted, extra-chance results are to be expected." These conclusions are exactly contrary to the facts that his own experiment, permitting more sources of error than any published ESP experiment, did *not* give extra-chance results, and that the consistency of error by the selected agent was precisely *not* illustrated by the results.

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

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

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

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

Criticism of Mathematical Evaluation

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

The common error of confusing two statistical tasks is apparent in the above suggestion. An analysis may be made of the data in order to *describe* it statistically. Fitting a curve to the experimental data does just this. Or a treatment of the data may attempt to evaluate its *difference* from a theoretical expectation, as is done in a chi-square test of goodness of fit. It has been assumed, and later shown empirically (5) that the binomial distribution is a good fit to chance expectation in these tests. If an extra-chance factor enters into the production of the experimental results there is no *a priori* reason why it should produce a normal, binomial, or any other distribution. We might just as well assume to begin with that the type of the distribution will be unknown. But since we know the type of chance expectation, and measurements *upon the assumption that the experimental results are chance deviants of that expectation* show that assumption to be highly improbable, then we can reasonably infer a non-chance factor. The experimenter may, of course, fit a curve to his experimental results, but information as to whether the results are chance or not must come from comparison with a chance hypothesis either theoretically or empirically based, and can be derived from no mathematical description of the test results.

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

Although the article purports to show the function of dependent probability in ESP data, no such demonstration is given; possibly because, in spite of the thesis, no reported ESP experiment has ever been conducted under the condition that the subject knows each card as it is guessed. And if Miss Becknell had consulted the reports of ESP research themselves, rather than limiting her bibliography to the remarks of critics, she might have found that the subject knows nothing about the order of the deck until after the 25 guesses are made.

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

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

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

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

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

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

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

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

Criticisms of Concepts and Implications

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

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

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

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

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

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

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

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

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

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

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

* * *

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

REFERENCES

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

5. GREENWOOD, J. A. Analysis of a large chance control series of ESP data. J. PARAPSYCHOL., **2**, 138-146.

6. GREENWOOD, J. A. An empirical investigation of some sampling problems. J. PARAPSYCHOL., 1938, **2**, 222-230.

7. GREENWOOD, J. A. Variance of the ESP call series. J. PARAPSYCHOL., 1938, **2**, 60-64.

8. GREENWOOD, J. A. & STUART, C. E. Mathematical techniques used in ESP research. J. PARAPSYCHOL., 1937, **1**, 206-225.

9. GREVILLE, T. N. E. Exact probabilities for the matching hypothesis. J. PARAPSYCHOL., 1938, **2**, 55-59.

10. HANSEN, F. C. C. & LEHMANN, A. Über unwillkürliches Flüstern. *Phil. Stud.*, 1895, **17**, 471-530.

11. HEINLEIN, C. P. & HEINLEIN, J. H. Critique of the premises and statistical methodology of parapsychology. *J. Psychol.*, 1938, **5**, 135-148.

12. HERR, D. L. A mathematical analysis of the experiments in extra-sensory perception. *J. Exper. Psychol.*, 1938, **22**, 491-495.

13. HUMPHREY, B. M. & CLARK, J. A. A comparison of clairvoyant and chance matching. J. PARAPSYCHOL., 1938, **2**, 31-37.

14. HUNTINGTON, E. V. A rating table for card matching experiments. J. PARAPSYCHOL., 1937, **1**, 292-294.

15. KELLOGG, C. E. Letters and notes. J. PARAPSYCHOL., 1938, **2**, 147-148.

16. KELLOGG, C. E. The problems of matching and sampling in the study of extra-sensory perception. *J. Abn. & Soc. Psychol.*, 1937, **32**, 462-479.

17. KELLOGG, C. E. The statistical techniques of ESP. *J. Gen. Psychol.*, 1938, **19**, 383-390.

18. KENNEDY, J. L. Suggestions concerning the nature and production of "extra-sensory" perception. [in manuscript].

19. KENNEDY, J. L. The visual cues from the backs of the ESP cards. *J. Psychol.*, 1938, **6**, 149-153.

20. LEUBA, C. An experiment to test the role of chance in ESP research. J. PARAPSYCHOL., 1938, **2**, 217-221.

21. RHINE, J. B. The question of sensory cues and the evidence. J. PARAPSYCHOL., 1937, **1**, 276-291.

22. RIESS, B. F. A case of high scores in card guessing at a distance. J. PARAPSYCHOL., 1937, **1**, 260-263.

23. ROGOSIN, H. Probability theory and extra-sensory perception. *J. Psychol.*, 1938, **5**, 265-270.

24. ROGOSIN, H. Some implications of extra-sensory perception. *Psychol. League Journ.*, 1938, **2**, 47-49.

25. STUART, C. E. A review of certain proposed hypotheses alternative to extra-sensory perception. *J. Abn. & Soc. Psychol.*, 1938, **33**, 57-70.

26. STUART, C. E. & GREENWOOD, J. A. Letters and notes. J. PARAPSYCHOL., 1938, **2**, 148-149.

27. STUART, C. E. & GREENWOOD, J. A. A review of criticisms of the mathematical evaluation of ESP data. J. PARAPSYCHOL., 1937, **1**, 295-304.

28. WARNER, L. A test case. J. PARAPSYCHOL., 1937, **1**, 234-238.

29. WILLOUGHBY, R. R. Prerequisites for a clairvoyance hypothesis. *J. Appl. Psychol.*, 1935, **19**, 543-550.