

Parapsychological Monographs

No. 10

On the Evaluation of  
Verbal Material in Parapsychology

J. G. PRATT

PARAPSYCHOLOGY FOUNDATION, INC.

29 West 57th Street, New York, N.Y. 10019

Biographical Note

# On the Evaluation of Verbal Material in Parapsychology

Dr. Pratt is well known for his contributions to parapsychology. He is author or co-author of approximately one hundred articles and books in the field. His most recent book is *Parapsychology: An Introduction* (Dutton, 1984).

## Biographical Note

*JOSEPH GAITHER PRATT, Ph.D., took his degree in psychology at Duke University in 1936 when the department was headed by Professor William McDougall. After working for two years in ESP research with Dr. Gardner Murphy at Columbia, Dr. Pratt became a member of the research staff of the Duke Parapsychology Laboratory. Since 1964 he has been doing full-time research in parapsychology in the Department of Psychiatry, University of Virginia, and he is now Assistant Professor there in the newly-established Division of Parapsychology.*

*Dr. Pratt is well known for his contributions to parapsychology, since he is author or co-author of approximately one hundred articles and books in this field. His most recent book is Parapsychology, An Insider's View of ESP (Dutton, 1966).*

Copyright © 1968 by Parapsychology Foundation, Inc.  
All Rights Reserved  
Library of Congress Catalog Card Number: 70-29884

ON THE EVALUATION OF  
VERBAL MATERIAL IN PARAPSYCHOLOGY

J. G. Pratt

Division of Parapsychology  
Department of Psychiatry  
University of Virginia

Parapsychology Foundation, Inc.  
29 West 57th Street, New York, N. Y. 10019

Copyright © 1969, by Parapsychology Foundation, Inc.

*All Rights Reserved*

Library of Congress Catalog Card Number: 70-94866

J4PN - 0-912328-14-2

ON THE EVALUATION OF  
VERBAL MATERIAL IN PARAPSYCHOLOGY

Department of Psychology  
University of Virginia

Manufactured in the United States of America

29 West 57th Street, New York, N. Y. 10019

*To Eileen J. Garrett*

whose remarkable talents and unquenchable thirst for personal answers and new scientific knowledge have led her to seek the investigators and to be sought by them.

## CONTENTS

I. Foreword . . . . .	9
II. Towards a Method of Evaluating Mediumistic Material . . . . .	15
III. Appraising Verbal Test Material in Parapsychology . . . . .	53
IV. Investigations Based on Verbal Material: Progress and Prospects . . . . .	67

# I

## FOREWORD

In 1934 the Boston Society for Psychic Research published J. B. Rhine's book *Extra-Sensory Perception*. In 1936 the same organization published, in its BULLETIN XXIII, my studies of Mrs. Eileen J. Garrett's mediumship under the title *Towards a Method of Evaluating Mediumistic Material*. Beyond sharing the same publisher, what else did these two works have in common?

The answer is: a great deal. They were both experimental studies done at Duke University and they were publications resulting from the new effort being made there to bring parapsychology into the mainstream of the strong movement toward quantitative experimental methods that was taking hold in American academic psychology at that time. The former publication presented ESP as an ability that is widely distributed in the general population and as one that lends itself to experimental investigation as do other normal abilities. The latter work showed through actual research that the non-quantitative material of "mental" mediumship can likewise be investigated by methods that meet the requirements of modern experimental science. Both studies owed a great debt to earlier efforts in their respective areas, and each made its contribution by adding to the accomplishments of previous investigations rather than by starting from the zero point.

In spite of these close similarities between the two publications when they appeared, the effects since that time of the two works are marked chiefly by contrasts. *Extra-Sensory Perception* was an immediate "hit." This is not to say that it was welcomed with acclaim on all sides, for there were probably more scientists who were bitterly critical of the book than there were scientists either ready to test the claims it set forth or willing to suspend judgment until others



had done so. But at least the book stirred up the kind of interest that guaranteed that the issues it raised would not be neglected and forgotten, and most of the developments in experimental parapsychology over the past three decades have, in one way or another, flowed from that publication.

*Towards a Method of Evaluating Mediumistic Material*, on the other hand, went virtually unnoticed, and the problem area with which it was concerned sank steadily into a deeper and deeper state of neglect. One might almost say that there was a seesaw effect: for the ESP end of the board to rise, it seemed necessary for the mediumship end to go down.

The eclipse of the latter as a research problem was nearly complete for a period of twelve years; then there appeared another publication (the Pratt-Birge method, presented in Part III of this monograph) that showed that the methodological problem posed by the phenomena of mediumship had not been completely abandoned. Like the earlier Boston Society Bulletin, this article was an effort to improve the methods available to the research worker in this area and it may have marked a turning point, for since it appeared in 1948 there has been a slow but steady reawakening of scientific interest among a few investigators.

Some of us have seen in this change welcome indications that research attention is once more turning to areas of parapsychology that had largely been ignored during the decades when statistical studies alone held sway. Now the change seems to be occurring as a genuine broadening of the field to include more of its legitimate areas of interest rather than as a swing away from experiments based upon restricted-choice, quantitative material. The need of parapsychology to push into new areas does not imply that there should be any abandonment of territory in which the most progress has recently been made. We must seek to advance on all fronts at the same time if the field is not to deserve after all the image of alternate periods of "seesaw" rises and falls that leave matters in the end too much as they were at the beginning.

The decision of the Parapsychology Foundation to republish my old Boston Society Bulletin is clearly related to the "New Look in Parapsychology," but one may still reasonably ask: In what way and to what degree? At a minimum level, this work was a serious effort

to improve research methodology, and as such it is a publication that many present-day workers in the field will wish to know. The demise of the Boston Society for Psychic Research shortly after the Bulletin was published and the fact that the limited issue has been out of print for about three decades makes the original (and only) edition as hard to locate as a Franklin car. Even workers who were already active in the field at that time often seem to have forgotten about this particular investigation<sup>1</sup> or they have been known to refer to it in such a way that newcomers to the field would not know where to turn to locate the original work.<sup>2</sup> Without laboring the point, two further instances can be mentioned (without citing names, to avoid causing embarrassment) to show that the lack of knowledge about this publication has been a stumbling block to writers in recent years. One Sunday supplement feature article that was ghost written under the name of an investigator in the field discussed this work at length, but went completely astray on the matter of whose work it was. Another investigator recently spent a lot of thought and time on "developing" the method that the Boston Bulletin had presented in 1936 and in writing a paper on what he thought was an original idea. So there is some need for workers to be informed about the history of their field, even if only to avoid the various kinds of mistakes they might otherwise make.

My own immediate acceptance of Mrs. Garrett's suggestion that the Parapsychology Foundation would like to reissue the 1936 Bulletin without doubt reflects a personal bias. I freely admit this fact, and I hope it will be understood and excused. After all, this Bulletin was my first publication in parapsychology, and I was pleased that it should be considered worthy of being rescued from oblivion.

But during the next few weeks I tested the idea by giving it more sober examination. If the work and expense of republication were justified, they would require better reasons than my personal satisfaction at seeing this old work once more in print. This re-examination confirmed the value of the plan but at the same time enlarged and reshaped the project.

---

<sup>1</sup> Rhine, L. E., "Parapsychology, then and now." *Journal of Parapsychology*, 1967, vol. 31, pp. 231-248.

<sup>2</sup> Birge, W. R., and Rhine, J. B., "Unusual types of persons tested for ESP: I. A professional medium." *Journal of Parapsychology*, 1942, vol. 6, pp. 85-94.

The Boston Society Bulletin was presented originally as only a step toward the kind of procedure required for the scientific evaluation of verbal mediumistic material. Further important steps have been taken since that time. To republish the 1936 study without taking account of these further developments would give an incorrect view of where we now stand on this issue in parapsychology.

This line of thinking led to the revised plan carried out in the following pages. The Boston Society Bulletin is reproduced, in abridged form, in Part II. Part III contains the main part of the article by W. R. Birge and myself, "Appraising Verbal Test Material in Parapsychology," published in 1948 in the *Journal of Parapsychology*. Because of the success of the Pratt-Birge method in meeting the major theoretical and practical deficiencies of the 1936 method, it was widely regarded among investigators as the standard evaluative procedure to be applied in working with verbal material.

This is not to say that the 1948 article was itself the last word in this area. Since that time two modifications of our method have been offered with the aim of making the statistical assessment more sensitive to the intrinsic quality of the verbal statements. (The details will be given later.) These and other considerations led to adding to the present publication a section which discusses developments that have taken place in parapsychology in those areas for which methods for the evaluation of verbal material are relevant.

One point to be considered in this concluding and updating section of the monograph will be whether the evidence from verbal material is ever so strong and so clear as to enable us safely to dispense with statistical evaluation. We need not only to have methods available for use where they are required, but we have the right as investigators to decide what method is best suited to the job at hand, including laying aside our statistical sieve if we have the good fortune to find unmistakable nuggets of fact.

This monograph is not only concerned with developments spanning the last three decades; it is oriented toward the research activities on the advancing frontier of parapsychology. Indeed, we are likely to find ourselves more involved in the years ahead with research calling for special methods of evaluation than we have been during the past three decades. The publication of this monograph must be justified in terms of its relevance to these real research needs.

To offset to a degree the additional length made necessary by the plan to do more than simply republish the 1936 Bulletin, some sections of that work have been omitted. These include a Foreword by Dr. Gardner Murphy, a concluding discussion by Dr. J. B. Rhine, and some of the appendices presenting mediumistic records to illustrate the method of analysis. For the student who may wish to refer to some of this material, copies of the original Bulletin can still be consulted in a few libraries.

I will end this introductory statement with a reference to the role that Mrs. Garrett played in this research beyond her indispensable participation as a subject. It was through her initiative that Mrs. Frances P. Bolton provided financial support for this research at Duke University. Mrs. Garrett then offered her services to Professor William McDougall, Chairman of the Department of Psychology at Duke, for a parapsychological study of her abilities.

Her first visit in 1934 led to a two-pronged investigation. One was a series of tests carried out under Dr. Rhine's direction in which Mrs. Garrett was tested for ESP with the standard card methods recently developed at Duke. The other was, of course, the study—presented in this monograph—of the abilities Mrs. Garrett showed in her mediumistic trance. The success achieved in both projects led to the second visit in 1935, and the two aspects of the work with her were described in separate publications.<sup>3</sup>

As a direct consequence of this research development, Mrs. Bolton was encouraged to support the work in this area at Duke University on a scale that made it possible formally to launch the Parapsychology Laboratory as a research center in the Psychology Department in the fall of 1935. The future historian of parapsychology should not overlook the fact that Mrs. Garrett was at that time active behind the scenes in promoting research in parapsychology, and not only in offering her talents as a sensitive. Her role in the developments of scientific parapsychology since that time will appear as a major one when the history of our field is finally written.

---

<sup>3</sup> The card test results were reported in: Rhine, J. B., "Telepathy and clairvoyance in the normal and trance states of a 'medium,'" *Character and Personality*, 1934, vol. 3, pp. 91-111.

## II

### TOWARDS A METHOD OF EVALUATING MEDIUMISTIC MATERIAL<sup>1</sup>

GIVEN: a number of records of a personal nature taken stenographically from a "medium" in trance, each record intended for a different subject. THE PROBLEM: to determine whether these records contain more correct information for the actual subjects than could be accounted for by mere random guessing on the part of the sensitive. Two steps were taken in this effort to answer this problem. In the first the subjects knew which were their own records. In the second, they did not know for whom each record was intended. All the subjects were asked to fit all the records to their individual cases. An estimation of the significance of the records was then arrived at by comparing the degree to which each description fitted the circumstances of the actual subject with the degree to which each applied to the remaining subjects.

This study is concerned with the *method* of evaluating mediumistic material, to see whether the correct information given is more than can be attributed to chance. The question of the ultimate source or means of acquiring knowledge without the known senses is not gone into.

Of these two questions concerning parapsychic capacities—*demonstration* and *explanation*—the latter is by far the more important, but the exclusion of chance as a hypothesis must be accomplished beyond a shadow of a reasonable doubt before the more important problem can properly be raised at all.

Many investigators of mediumship have been convinced of the significance of mediumistic records, depending almost entirely upon the subjective judgment of the sitter. Some of the conclusions arrived at in this manner have been fully justified, but investigators must

---

<sup>1</sup> Reprinted, with minor adaptations and with omissions indicated, from *Bulletin* XXXIII, Boston Society for Psychic Research, March, 1936.

recognize that the nature of this subject is such that they can afford to spare no pains in striving for more objective methods. If a sitter is justified in accepting his own record as significant, there should be some way of evaluating it which would be convincing to a larger number of fair-minded readers.

But why this concern at the present time about the method of evaluating mediumistic material? The study of mediumship is important for parapsychology only if it is possible to show whether or not the sensitives have parapsychic capacities. This means that if trance utterances are to be studied, there must be evolved some method of determining whether the sensitives show more correct information than can be attributed to other factors. As these trance descriptions are often of a peculiarly general nature and very difficult to judge, special attention to the method of evaluation may be required. In the treatment of such material, this is really the prior question. Sometimes it has been dealt with in one way, sometimes in another; many times it has been almost entirely overlooked, the problem seeming to disappear because of the impressiveness of the records.

But there is a very large body of records of a highly personal and descriptive nature in which a little that is correct seems to be intermingled with much that is incorrect. At the same time, some of the correct statements are sometimes of a highly general nature, so that it is impossible to judge their significance off-hand. With records of this kind, there is great likelihood that such prejudices as the "will to believe" or the "will to disbelieve" will influence one's judgments, and it is here that emphasis upon objective methods of evaluation is necessary. The records with which this study is concerned are of this personally descriptive variety.

Strong scientific interest in mediumistic material has been shown for a long time, chiefly since the beginning of the studies of Mrs. Piper in the last part of the nineteenth century. Of course, interest in what information a sensitive gives implies an evaluation, and this necessarily indicates the use of some kind of method. Perhaps the most usual procedure has been to depend entirely upon the judgment of the subject. In the use of this method, the subject either is actually present at the experiment or receives the record with the full knowledge that it is intended to apply to himself. He annotates

the record, or says outright whether it is good or poor. There is no attempt to fit the record to other specific individuals. Where there is vagueness and ambiguity in the material and where, in addition, there may be a strong motive for "communicating," dependence to this full extent upon the judgment of the subject is short of ideal, to say the least.

But recently there has been felt in certain circles a need for a more objective approach to such material. There are two studies which make most important contributions to the methods of evaluating such descriptive material, and a third which develops a mathematical formula for this purpose. The first two of these are the studies by Mr. H. F. Saltmarsh<sup>2</sup> and by Dr. J. F. Thomas.<sup>3</sup>

In the study by Mr. Saltmarsh, excellent safeguards were taken against the medium's use of sensory cues or rational inference of any preventable kind. The subjects who were permitted to be present at experiments were, with two exceptions, anonymous. Some experiments were conducted with subjects absent. For these "absent sitter" experiments, relics were used to aid the sensitive, and the arrangements were such that no one knew, by any normal means, for whom each experiment was being conducted.

Mr. Saltmarsh's procedure in the evaluation of results was to submit the analyzed material for annotation to actual sitters and to non-sitters. The actual sitters knew that the material being annotated by them was intended to apply to their circumstances, while the reverse would necessarily be true for non-sitters. Here we have the factor of the knowledge that the material was or was not supposed to apply to the scorer. As it is useless to argue whether this factor is strong or weak, the only adequate procedure is to eliminate it. This might well have been done in Mr. Saltmarsh's study.

Mr. Saltmarsh analyzed the records into items, and these were classified into three groups with different values assigned arbitrarily to each. *Clichés* received a value of 1, definite statements a value of 5, and characteristic statements 20. Mr. Saltmarsh served as the

---

<sup>2</sup> Saltmarsh, H. F., "Report on the investigation of some sittings with Mrs. Warren Elliott," *Proceedings, Society for Psychical Research*, 1930-31, vol. 39, pp. 47-184.

<sup>3</sup> Thomas, J. F., "A study in the statistical evaluation of the mental content of certain trance phenomena." Ph.D. Thesis, Duke University Library, 1933.

judge of whether each item was true or false, using as his basis the notes made by subjects and non-subjects. A determination of the anti-chance value was made; in this the average value of sitters' scores was more than eight times the average value of non-sitters' scores. This was taken to indicate a real knowledge-content in the medium's material.

Dr. Thomas's study involved a number of records secured with different sensitives. He was always the subject. Some experiments were conducted in his presence, some not. In all experiments precautions were taken to conceal his identity from the sensitive.

In the actual evaluation, three general procedures were followed.

(1) Many of the sittings were analyzed by a topic-and-point procedure, and checked by Dr. Thomas to see how they fitted his case. The results were stated in terms of per cents applicable, incorrect, answers unknown, and irrelevant. Usually, around ninety per cent were found to be correct for Dr. Thomas.

(2) A questionnaire form was made from some of his records. This was checked for self-applicability by a number of people, including Dr. Thomas. The extent to which these items were true for all the people answering the questionnaire was thereby determined, and the case of Dr. Thomas could be given a more or less definite anti-chance value by the way in which he stood out from the group. Stated mathematically, a frequency distribution was made of all the scores on the questionnaire, and the significance of Dr. Thomas's score was determined in terms of the standard error and his departure from the mean score.

(3) In the third procedure, points from his records were presented to two committees, to have them estimate the general proportion of people for whom these items would be true. Dr. Thomas, again, checked these same items for himself. The averages of the committees' estimates were used, on the basis of Dr. Thomas's checking of the same items for himself, in the Saltmarsh-Soal formulas. (These are the formulas proposed in the third study, which we examine next.) These formulas provided a means of determining the odds against this material's applying to anyone by mere chance to the same degree as it did to Dr. Thomas.

By the procedures used, the records evaluated in his study proved



to be overwhelmingly significant of the possession of knowledge by the sensitives which random guessing could not have given.

While the two contributions to methods of evaluating mediumistic records which we have examined have many merits, they still allowed the personal factor to enter into the judgment of how well the material fitted the circumstances of the actual subject. The third study to be examined<sup>4</sup> develops mathematical formulas by which such records may be evaluated, and Mr. Saltmarsh makes a specimen application of the method to a record from the Mrs. Warren Elliott series.

The proposed method consists of formulas developed by Messrs. Soal and Fisher for determining the probability that a certain series of statements will, through chance alone, apply as well as they actually do to a particular individual, when the way in which they apply to a larger general group is known. However, they call for the use of definite values which must in some way be derived from the records as they are taken from the sensitive. The results of applying the formulas can be no more trustworthy than the figures used in the application. In describing his proposed method, Mr. Saltmarsh gives a specimen application of the formulas. This is precisely what is attempted in a broader scope and in a more detailed fashion in the present study. I will show first how Mr. Saltmarsh makes this transition from the record to be evaluated to the formulas, and then we can proceed at once to a description of how this step is taken in the present study.

Mr. Saltmarsh analyzed a short record into points. These were then submitted to the actual subject for his notes. Presumably, the subject knew that this was his own record. Mr. Saltmarsh decided on the basis of the notes whether the items were true or false for the subject. He further estimated for what proportion of people in general these same items would be true. Thus he arrived at the values needed for the application of the formulas to the record.

In his specimen use of the method, Mr. Saltmarsh recognized fully that the determination of these general probability ratios should

---

<sup>4</sup> Saltmarsh, H. F., and Soal, S. G., "A method for estimating the supernormal content of mediumistic communications," *Proceedings, S.P.R.*, 1929-1931, vol. 39, pp. 266-271. These formulas, which have the additional authority of Prof. R. A. Fisher, are not reproduced here because they are so much more clearly presented in the original report than could be done in shorter compass.

not rest solely upon the judgment of one individual. Instead, he proposes a committee procedure, such as that later employed by Dr. Thomas. But in none of the previous studies have the procedures taken fully into account the subjective factor in the actual subject's marking of the material for himself.

We have seen that in his specimen use of the method, Mr. Saltmarsh had first to get answers to two questions: (1) Are the statements of the sensitive correct for the actual subject? and (2) For what part of the population in general are these same statements true? Now, obviously, the first question must generally be answered by the subject himself. But it is clearly important to have his answers as unbiased as possible. Obviously, also, it is not possible to interview an unlimited number of people to get a perfect answer to the second question.

From the brief review which has been made, it is clear that important steps toward an adequate objective method of dealing with trance material have been made. But it is also clear that further steps in this direction remain to be taken. These are the elimination of the personal factor in the actual subject's estimation of how well descriptions fit his own circumstances and some procedure for arriving at more dependable ratios of general applicability for the points of the descriptions. These two steps are attempted in the present study.

The sensitive employed in this study was the English medium, Mrs. Eileen Garrett. She made two visits to Duke University, each one being for a period of about two weeks. The first two weeks of experimentation occurred in the middle of April, 1934. More records were taken during this period than could conveniently be studied; so they were arbitrarily limited in number by taking the first record of each subject who sat with Mrs. Garrett. As there were twelve subjects, this provided twelve records from the series of April, 1934, to be evaluated. These are referred to throughout the remainder of this study as Series I. The second period of experimentation fell in the latter part of January, 1935, and only one record was taken for each of fifteen subjects. The whole number was treated. These are designated Series II.

During both periods of experimentation the sensitive was being subjected to a strenuous program of research in which she was also the subject of an investigation (in both normal and trance states)

of her parapsychic capacities, which work has already been reported.<sup>5</sup> The sensitive also gave a few sittings privately for individuals during both periods. These did not come within the scope of the experimental material reported here. During the first period, the sensitive performed from one to three trance experiments a day for a period of fourteen days. The trance work of the second series covered about the same number of days, with one and very rarely two experiments a day.

The experiments were conducted at Duke University and were a part of the program of parapsychological investigations being carried out there. All the utterances of the sensitive and everything else said during an experiment were recorded *verbatim* by a stenographer. As precautions against the sensitive's use of knowledge acquired through her senses, the names of the subjects were kept in the strictest secrecy and nothing was spoken by the subjects of Series I in the presence of the sensitive.

Subjects for these experiments, selected from members of the university community, were invited privately to participate. "Suitable" subjects were people who did not depart radically from the normal circumstances of the community life, who were thought likely to have an interest of one kind or another in the experiments, and who in the main were known to have deceased connections with whom they might presumably like to communicate. The ages of subjects ranged from twenty-two to seventy, with the majority at or below middle age. In Series I there were five males and seven females; in Series II, five males and ten females. Adherence to no particular hypothesis was required, but subjects were asked for technical reasons to assume at the time of the experiment an attitude as sympathetic as possible to the survival hypothesis. All were told that the study would involve a drain upon their time and all accepted the invitation with this understanding. Subjects were promised that their names would not be published. All experiments were conducted with strict privacy. Especial care was taken to guard the identity of the sitter. In Series II only two persons, Dr. Rhine and myself, knew of the selection made daily, and these choices were commonly made less than a day ahead.

Mrs. Garrett seems to be able to go into and come out of

---

<sup>5</sup> Rhine, J. B., "Telepathy and clairvoyance in the normal and trance states of a 'medium,'" *Character and Personality*, 1934, vol. III, pp. 91-111.

trance at will; at least, she never failed while under the observation of those connected with these investigations. To bring on the trance, she sits in a soft chair, slumps down or "sprawls" with her feet far forward and her head thrown back. She appears to relax completely. Her eyes roll upward in their sockets. After a few moments her breath begins to come in short gasps, and she utters low moans. The eyes are by this time closed, though during the trance the lids seem to flicker continuously. In a short while she crosses her arms across her breast and begins to sway back and forth in her chair. The voice of the control is then heard; it has a low, masculine quality. His first utterance seldom varies from the stereotyped phrases: "It is I, Uvani. I bring you greetings, friends." Uvani claims to be the spirit of a deceased Arab. He is very co-operative in experimentation.<sup>6</sup> Upon being told that a subject is ready in the same room with himself or in the next room with the intervening door closed, he begins his remarks for that subject and continues until he talks himself out. He never tries, as far as I have observed, to get leads from the subject or those present. Sometimes he apologizes because he does not feel that he is having much success. He announces when he is finished, and asks if there is anything more desired from him for this time. Two or three long experiments may be conducted during one period of trance, consuming from one to two hours. When Uvani is dismissed, Mrs. Garrett arouses herself by stretching and yawning as though coming out of a natural sleep. She appears to be fully normal in about a minute's time.

The actual laboratory procedures for the two series differed in one important respect. During Series I the subjects were present in the same room as the sensitive and heard her give the descriptions which were supposed to apply to them. In Series II, however, they sat in an adjoining room with a closed door between and a noisy electric fan running. They neither heard what the sensitive said nor were informed which was their record until the evaluation of the records was completed. In both series there were present at each trance experiment a stenographer and some one in charge. The

---

<sup>6</sup> The practice followed in these experiments was to devote a trance at the first of each series to explaining the nature of the experiments to Uvani and asking his co-operation. This he readily gave. Mrs. Garrett in her waking state professes to be committed to no particular theory of the nature of her trance state. She is quite favorable to—and in fact welcomes—scientific investigation.

subjects were asked to bring tokens which were objects of affection to some deceased relative or friend. These were not always used, however; sometimes they were given to Uvani by the experimenter when the "control" appeared to have talked himself out. Hints and aids to the sensitive were studiously avoided.

In Series I the subjects were introduced to the presence of the sensitive after she had assumed the trance state. They were seated behind the sensitive and did not speak while in her presence. Subjects left before Mrs. Garrett resumed her normal state. When this body of material (Series I) was submitted later to each subject, it was possible that part of it might be recognized as his own record. In this respect Series I failed to control the subjective factors in the subjects' judgments of how well the material applied to themselves. For this reason, Series II is of much more importance in this study.

I analyzed the records into items. They were then *all* submitted to *all* the subjects to apply to their individual circumstances. That is to say, each subject was asked to check not only his own material but also that of every other subject *as though* it were his own. In Series II this took the form of checking a number of records for self-applicability without knowing which was one's own. From these data the values substitutable in the formulas were derived in a manner to be described below.

In analyzing a record into items for checking by the subjects, the aim, of course, was to stay close to the meaning of the sensitive. The analyzing was done by myself. In some cases the material may have been ambiguous and my interpretation therefore erroneous. But as I knew very little of the circumstances of blood relationship of the subjects, my few cases of faulty interpretation could hardly have affected the outcome. A complete record from Series II, both as it was taken by the stenographer and as it was analyzed for submitting to the subjects, is given in Appendix A, pp. 33-44. The subjects in this series saw only the analyzed forms. This particular record was a "good" one, being that of the fourteenth subject in Table II (p. 29). It was taken near the middle of the series, was longer than the average, and offered more than average difficulty of analysis.

Each subject was assigned some arbitrary symbol by the secretary to distinguish him. The secretary did not provide me with a key to

the subjects until I had completed the analysis, and in most cases I could not have consciously identified the subject. The analyzed forms were numbered in sections in the order in which they were submitted to the subjects for checking. All this applies only to Series II, from which the record given in the Appendix is taken.

The sample record shown also reveals that not all that was said by the sensitive was included in the analyzed form. The rule followed was to omit all direct references of a personal, descriptive kind to the subject himself. The sensitive always devotes her first remarks to the sitter, commenting upon his health and speaking generally about his personality, usually in a complimentary fashion. This part of each record was always easy to identify and could be omitted with fairness to the whole. Doing this in both series of the present study made it possible to evaluate the records of more subjects.

The records of Series I were embodied in a single mimeographed form. This form contained the itemized material based upon the records of the first experiments of twelve subjects. The complete form used in Series I is given in Appendix B [omitted from this monograph].

The method outlined here did not aim at an exact determination of the individual "hits" made by the sensitive. To do this for descriptive material of this personal kind is a great difficulty, if not an impossibility. The aim was, however, to determine whether a series of statements intended for one person had a significantly greater weight of truth for that person than for a number of others of similar general circumstances. For this purpose it was not necessary that all be allowed unlimited scope in attempting to apply items to themselves. What *was* important was that the same arbitrary rules be imposed upon all. The purpose of the first part of the form used in Series I was to define arbitrarily the scope within which those taking part should feel free to apply each item to themselves. . . .

. . . A much less elaborate set of instructions (Appendices C and D, pp. 45-48) was used in Series II. This was felt to be possible because eight subjects were common to the two series and because in Series II each analyzed record was being sent out as a separate installment.

The form in Series I was mailed to the twelve subjects and to twenty-five non-subjects, the latter having been asked by mail

whether they would co-operate by checking the form for themselves. Only fifteen of the twenty-five returned the checked form, however; these and the twelve subjects are all that enter into the calculations of Series I. The purpose in using non-subjects was to secure more representative ratios of general applicability for the individual items of the descriptions given by the sensitive, by having more people apply them to themselves. . . .

. . . The analyzed records of Series I were presented in the form in unbroken succession. The order of subjects was a random one. Extrinsic hints of the persons for whom the sensitive intended various parts of the material were carefully avoided. It was thought that subjects might fail, under these conditions, to recognize which material had been intended for them. . . . Each item could be answered in one of four ways: checked, crossed, question-marked, or left blank. Only those which the actual subject answered for himself by checking or crossing (true or false) entered into the calculations.

The plan of organizing the points for checking was chosen from several alternatives as particularly suitable to provide data for treatment by the Saltmarsh-Soal method. Whenever the sensitive introduced a new personality—usually two or three are prominent in each record, separate accounts being woven about each—an item of a highly general nature was used in the analysis of the material to mark this point. Detailed items of description were brought in under these opening items. In the opening items were included the indicated degree of relationship (if any were given), whether the person were living or deceased, and sex. I expected that most of the opening items would be correct for every answerer. Where everyone checked as correct one of these items, the answers bore, in the method, neither for nor against a favorable result. In such a case the whole matter hinged on the items of detailed description. Having to answer “no” to an opening item was taken to be indicative of a total “miss” on the part of the sensitive. In this case the other items had to be skipped as having no bearing for the answerer. But it seemed appropriate to give some weight to the fact of the correctness or incorrectness of the sensitive’s *having introduced a given personality into the record at all*. This the general opening items served to do. For example, suppose an opening item to be inapplicable to the actual subject. The fact of the sensitive’s having erroneously spoken of a given

relative of the actual subject weighed *against* her demonstration of extra-sensory knowledge. The amount of this negative weight would, of course, depend upon the extent to which the same general item proved to be true for the remaining answerers.

For both series, the data of personal applicability of each record were used in the following manner. A chart was made in which all the answers of all subjects to all items could be recorded, that of the actual subject being in the right-hand column. When this was filled in, the answers of all other participants for particular items were given in the rows across the chart. Those items which the actual subject checked were the ones for which the ratios of general probability were positive; those which the real subject crossed, negative. Only the items scored by the actual subject as true or false were used in the final evaluation. The ratios of probability for each item were obtained from the responses of all the answerers except the actual subject. For each item this ratio was the number answering yes over the combined number answering yes and no. These ratios of probability were thus actually found by having a number of people fit the records to their own circumstances.

The ratios were converted into the nearest two-place decimals. The necessary values for each item could then be substituted in the formulas directly from the table published in Appendix E, pp. 48-51. The use that was made of this table is explained in the same Appendix. Those who wish to pursue the matter further may consult the original Saltmarsh-Soal report.

It was necessary to adopt a rule for the treatment of those items for which none of the answers, except that of the actual subject, was yes. Had the value of the probability ratio actually found been used—a zero value—one such item alone would have been “infinitely” significant. For this would be tantamount to saying that the actual subject was the only person living for whom this item was true. In such cases the procedure followed was to assume that had one more person answered this item, his response would have been a check and that the numerator in the ratio would thereby have become 1. This plan operated always, insofar as it departed from the facts of the case, in the direction of reducing the significance of the experimental records. It seemed, therefore, a safe thing to do.



The results of the evaluation of the records in Series I by this procedure are given in Table I. The subjects are simply numbered from 1 to 12 in the order in which their records occurred in the questionnaire form. The real score of each experiment, determined as described, is given in the second column. Some of these real scores are quite small, indicating that the sensitive either gave only a short account or that she was mistaken in her statements for the actual subject in about the same degree as she was correct. The standard deviation of each experiment is given in the third column; this is the measure of how much the real value might be expected to vary by chance. In the fourth column occurs the result of dividing the real score of the record by the standard deviation, a measure known as the critical ratio. Those records which were more wrong than right for the actual subject have minus signs before the real-score-of-record column entries and minus signs in parentheses after the value column entries.

A critical ratio of 2.5 or 3.0 is usually accepted as large enough to exclude chance factors as accounting for the result. The odds against chance for these two values of the critical ratio are 80 to 1 and 370 to 1, respectively.

TABLE I  
*Results of the Application of the Saltmarsh-Soal  
Formulas to Series I*

Subject	Real	Standard Deviation	Value of Record:
	Score of Record		Critical Ratio
1	.0017	.9024	.00
2	-1.4276	1.0625	1.35(-)
3	.6923	.9507	.73
4	-1.4532	.8077	1.80(-)
5	.2112	1.4683	.14
6	2.6672	.7437	3.60
7	-.6885	1.7561	.45(-)
8	.9134	1.2767	.71
9	3.1071	1.4000	2.22
10	2.0777	.8196	2.53
11	2.7863	1.3297	2.10
12	4.3934	1.3882	3.16
Total	13.2800	4.1300	3.22

Table I shows that if we accept a value of at least three times the standard deviation as statistically significant, two of the twelve records treated must be so considered. Three more show critical ratios of more than 2. Of the three records which gave a negative value, none approaches significance. Considering the whole body of material as a unit, the critical ratio is 3.22. The odds against the occurrence of such a value by chance alone are about 725 to 1. This effectively excludes chance as the explanation. Is this result to be accepted, therefore, as a statistical demonstration of real knowledge-content in these records which the sensitive manifested without the use of the channels of her senses? Or is there some loophole in the procedure?

One opportunity for spurious results in Series I has already been mentioned: the subjects heard their records taken and thus had direct knowledge of which records were intended to apply to them. Series II was planned from the outset in a manner to eliminate this personal factor. This was accomplished by the procedure, already described, of conducting the experiments with the subjects out of sight and out of hearing. Subjects marked records without knowing whose they were applying to themselves at the time.

Fifteen subjects served in Series II, and each subject later fitted all the material to his own circumstances. No additional answerers were employed. Upon making appointments with the subjects shortly before their experiments, I instructed them to come to a particular room by a route which would not take them by the room where Mrs. Garrett was. Just before going into the subjects' room for his experiment, each subject was given the instructions to read which are given in Appendix C, pp. 45-46.

In this series each record was analyzed (by myself) as a unit and sent out by mail separately, shortly after the completion of the experiments, at intervals of one to three days. The introductory statement given in Appendix D, pp. 46-48, was sent to each subject with the first analyzed record.

With two exceptions, all fifteen reports were returned for all experiments. As the actual subjects submitted their markings of their own experiments in every case, hearing in two cases from only fourteen subjects did not interfere appreciably with the evaluation. The data were treated in Series II in the manner described for the earlier experiments.

The statistical results of Series II are given in Table II.

TABLE II  
*Results of Series II*

Subject	Real Score of Record	Standard Deviation	Value of Record: Critical Ratio
1	.3780	1.2970	.291
2	1.6559	1.5649	1.06
3	-.4721	2.0732	.288(-)
4	3.8165	1.0436	3.66
5	11.9139	1.6006	7.44
6	1.0314	1.4053	.73
7	-.6015	1.4128	.30(-)
8	-.3108	1.3168	.12(-)
9	-.5731	.8814	.64(-)
10	-.1588	.1962	.81(-)
11	1.3216	1.3416	.99
12	-.2516	.8970	.28(-)
13	.0461	.3410	.14
14	5.4295	1.3660	3.04
<sup>1</sup> 12*	1.9818	1.1983	1.58
Total	25.2068	4.9650	5.10

\* This represents a reevaluation of the record of subject 12 of Series I. The significance of using this record again is explained later.

The results of Series II are not greatly different from those found earlier. But in this case the total series gives odds of about 1,700,000 to 1 against chance as a sufficient explanation.

I have said before that this study is not concerned with the isolation of striking "hits" as much as with developing an adequate method of dealing with a body of material as a whole. In the present case, the body of material really consists of all the experiments in each series. The result of each series as a whole is the most important thing for consideration, particularly that of Series II, where the personal factor was eliminated. This does not mean, however, that we may not examine the results of individual records in order to see how they severally contribute to the general result.

Restricting our attention to Series II and considering this as fifteen individual experiments, we find that in three cases the critical

ratios are more than 3.0. This is true in spite of the fact that none of these subjects knew that the records which appeared to be particularly applicable to them were actually intended so to be by the sensitive. The control of the factor of personal prejudice in answers seems here to have been adequate. But may there not still be some manner by which a seemingly significant result might have come about spuriously? May not some general peculiarity in the manner of answering of a particular subject have led to a distortion in the result?

One thing immediately suggests itself as such a possibility. This is a general tendency toward the exaggeration of how well the material fits. Or there may be a particularly broad scope of relations and circumstances on which an answerer might draw for applicability. If a minority of people scoring the material actually manifested such traits, their particular records might appear individually to be significant, as a result of the contrast between their answers and the larger body containing more negative responses. However, such a "yes" tendency controls itself as far as the whole series is concerned. The marked gain to the individual records of the "yes" scorer is offset by slight losses to all the others, due to this same person's affirmative answers for that material.

It would, however, be important to know whether any weight can be attached to the individual records which appear significant in themselves. Before we can do this, some control of the "yes" tendency of these subjects must be made.

Such a control was made effective in the case of the fourteen subjects whose results are reported in Series II. This consisted in counting the number of yes and no answers made by each subject in material *not his own* and taking the percentage that the yes answers were of this total number. In a similar way the percentage was obtained for the subject's *own record*. By getting these percentages for all subjects, an average of the relation of yes to no answers was found. The percentages for each subject could then be compared with this average. Suppose that a subject made about the same percentage of yes answers in both his own and other material and that this average was not far from the average of the group. This would indicate nothing more than that the record of this subject fitted his circumstances only as well as would be expected by chance. Imagine a case, however, in which his percentage of yes answers for

his own and other material were equal and *much above* the general average for the group. Here a general tendency to answer yes would be indicated. In such a case a highly significant result for the individual experiment would very probably be misleading as far as a demonstration of extra-sensory knowledge is concerned. The reader will recall that in the treatment of the data an individual record received a good score in proportion to the extent to which the items were correct for the actual subject more than for the remaining participants. But any record would appear to be true for a person who gives a high percentage of yes answers to *all* records. But suppose, again, that the subject's percentage of yes answers for his own material is much higher than the general average, while his percentage of yeses for other records is at or around the general group average. Here no general tendency to answer yes would be shown, and a significant result could not be accounted for on this ground.

Carrying out this procedure, it was found that, in general, the positive answers were forty-five per cent of the combined number of positive and negative ones. The following table shows the percentages for the three subjects showing significant records in Series II. It is of course the actual application of this control to the best records which is of any concern.

TABLE III

*Showing the Percentage of Positive Answers to the Total Number of Positive and Negative for the Three Outstanding Records in Series II*

Subject	Own Material	Other Material
4	72	47
5	65	72
14	61	47

Subject 5 stands out immediately as a person showing a general tendency to give positive answers. The other two subjects show no such tendency. So it looks as though the first attempt to apply this proposed method for the quantitative evaluation of trance material has produced two individual experiments which may be accepted, by statistical standards, as significant of something other than those factors brought under control by the procedure.

While fifteen subjects participated in Series II, one simply made a note on his own record that nothing in it applied to his case, and

thereby necessitated leaving it out of consideration. In its place, a record which showed a favorable result in Series I was introduced without the knowledge of the actual subject (who was serving also in Series II). The object was to ascertain how two scorings of the same material at six months' interval would compare. This subject later reported that he failed to recognize the material as his own of the previous series. The subject gave the same answers as before to eighty-nine per cent of the items. This indicates that in the case of this subject, at least, the answers were being determined by fitting the descriptions to the same circumstances.

The result of the evaluation of this record in Series I gave a critical ratio of 3.16. The odds against finding a record which fits as well as this by chance alone are 600 to 1. The result of the re-evaluation (Subject <sup>1</sup>12 in Table II) gives a critical ration of only 1.58, with odds of 8 to 1 against chance. That is to say, in the second series, this record appeared to fall short of significance. Since the manner in which the subject scored the record corresponded so closely in each case, this discrepancy in results appears rather surprising. There seems to be only one possible explanation, namely: differences in the ratios of general probability which were secured by the remaining answerers in the two series. But the fact that Series I with twenty-four answerers contributing to these ratios gave a higher score for this record than Series II, where only fourteen answerers contributed, rather strengthens the significant result which was found for Series II as a whole. This suggests that the subjects in Series II were more conservative answerers than those in Series I; this gives rise to the speculation that if the result of Series II varied from the true result because of the small number of subjects, it did so on the safe side. It speaks well for the method suggested here that the smaller number of participants gave the more striking result. But it suggests the relative character of determinations of this kind even at best, as the problem stands today.

In view of the fact that some readers may feel inclined to judge the capacities of the sensitive on the basis of this work, two points may be mentioned here again. First is the fact that she was required to give a large number of sittings to many people within a short period of time. At the same time she was working as a subject for other experiments. The other consideration is that of the personnel

of the subjects employed as "sitters": with two exceptions (both of them in Series I) they were inexperienced in experiments of this nature. The surprising feature has been that the results are as good as this study has shown them to be.

If this method is to have wide application in studies in mediumistic material in the future, it may have to be adapted to treating bodies of material secured by a few "good" subjects. It is doubted whether this adaptation can be made without some loss of assurance for the reliability of the results. But some features of this study are immediately available for all studies which involve the evaluation of mediumistic trance utterances. The most important is the control of the subjective factor in the evaluation of the material.

The method herein proposed might take some other form than that which it has assumed in this particular study. Its essential aims beyond those of previous methods are (1) to secure the judgments of applicability of the subject for his own material without any opportunity for personal prejudice to play a part, and (2) to arrive at ratios of general applicability on as nearly an empirical basis as possible. If we are going to claim veridicality beyond what we might expect by chance for descriptive records of the kind treated in this study, these aims seem essential, by whatever method they be achieved.

## APPENDICES

### APPENDIX A

A sample of a complete record from Series II as it was taken stenographically (Subject 14, Table II) is printed below. In the right-hand column is given the analyzed form which was sent out to the subjects. This offers an opportunity for observing what was done to a record when it was analyzed, as well as illustrating the nature of the personal material which was omitted from the evaluation.

[Mrs. Garrett usually begins her remarks as "Uvani" in trance by referring personally to some conditions of the subject's own character, personality, and general state of health. In the analysis of the records I adopted the rule that all direct references of a personal nature to the subjects themselves would be omitted. These were thought to be particularly hard to evaluate, and they could perfectly

well be omitted consistently from all experiments. It will be seen that objective descriptions of the subject's environment and behavior are included, e.g., the description of the subject's room and writing activity given below.]

#### SUBJECT 14

It is I, Uvani. I give you greetings, friends. Peace be with you. (Thank you, Uvani. I have here again an object, Uvani, which I hope will be able to help you in making contact.) It is interesting that even as you spoke of that the communicant might have been thinking of giving to you a little chain with keys that were evidently in the communicant's mind before, because that is the thing that is uppermost. May I speak to you a little of the communicant? Here again I am getting quite a distinctive touch of ability. There is ability to endeavor, concentration, organization, and to study, and there is a great deal of ability in the communicant's make-up to overcome very great difficulties. Just at the moment I have to say that, please, and stress it a little because it would seem that the communicant is going through a process, and that process is a rather depressing one. Indeed it might seem that in the last three or four days the communicant has felt a little depressed, a little unable to see a very clear process ahead, and there had been as near as possible a giving up of things. The communicant may not have put these into words but the giving up has been very much in the mind, and I think it would be a very great pity, since the only thing that may happen is a retiring of a goal but by no matter of means the missing of that goal. I am speaking of somebody acutely sensitive, who has a good deal of gaiety in make-up and a great deal of gravity—a very psychic personality whose conscientious regard for the truth of things is perhaps uppermost at all times. The communicant has no small legal ability as well as a mind that can deal with delicacies, and I would not be surprised if that did not transpire with a good deal of writing to be done. Already there has been a good deal of writing but circumstances have put that aside; nevertheless it has its accomplishments. The communicant is one who will go very great length in desires closest to heart and at same time be useful to many in a legal attitude, to many strata, to many sides; for instance in literary associations, and is an organizer in things always helpful. The psychic quality is particularly strong, and by no means without ability to help if the subject were a mathematical one. You do see the point? (Yes.) A lack of little financial conditions at the moment may make difficulties, but on the other hand there is such an attitude of ability to get over this difficulty as we speak



of that this difficulty is almost ceased to be one. I want to go for a little moment to the communicant's room and to say that the room has an austerity in keeping rather with the general attitude of the communicant. There is a simplicity and a great love of growing things, of green and the like.

There is, on a table on the right-hand side of the room, please, a strange little box that may indeed have emanated in Germany or Saxony. It does not seem to be a work of the present day; it looks as though it may have been done many years ago, oh, perhaps eighty years ago. It is like hand carving. That box does contain little mementos, but nothing I can very easily distinguish. It is unique in pattern.

There would seem to be a picture in that room also and that picture is not hanging up and in the mind of the communicant it might be placed for all to see. It is a picture of a man, clean-shaven, and about the age 56 or 58, very fair, blue eyed, and a man of very great spiritual quality. That man would give to me the impression of being the father of the communicant. I have no impression that he has passed over. I only know the picture is valued very much and that it is something that always accompanies the communicant. The man of the picture seems to give to me the impression that in the last five years he has not been in very good health. Two years ago he was having a very difficult time. He has been

1. There is a "strange little box" of some kind in your room. (If not, skip to Item 10.)
2. It is on a table.
3. On one side as you enter.
4. The box is German or Saxon in make.
5. It is not modern.
6. Made about 80 years ago.
7. Hand carved.
8. It contains mementos.
9. Unique in pattern.
10. There is a picture of a man in the room which may or may not be hanging up. (If not, skip to Item 24.)
11. The man in the picture is clean-shaven.
12. His age is 54-60.
13. Very fair complexioned.
14. Blue eyed.
15. He is a man "of very great spiritual quality."
16. He is your father.
17. Living.
18. The picture is valued very highly.
19. You always keep it with you wherever you live.
20. The man of the picture has not been in good health for the last five years.
21. He was particularly ill two years ago.

better since but still is not what you would call exactly robust.

The difficulty of health would lie in the area of the kidneys, and I think that at certain seasons of the year—particularly around this time of the year and October—that this man of the picture sometimes has quite a bad time, and such a bad time two years ago.

I want to go now from the little box and the picture which seem to be on the left hand side rather like on the bureau or a table as you enter the room. Now I want you to look on the right hand side for a moment with me and I have a feeling I am coming into contact with some books. It would seem the room had been changed around, things are not where they had been before, so this personality of the communicant is one who very often likes to make changes in things that are around him, if not in his own surroundings. And by the way, when I say "him" I mean only figuratively in this case. I want to draw the attention here to like a book. It is a book in which the communicant has been writing. There has been a good deal of doubt as to the ability to make of this writing something intensely fine. Hence the book has been put in a drawer unfinished. When I call it a book, I may more clearly state to you it is like a manuscript not yet bound but many leaves,

22. He has kidney trouble.

23. His health suffers most in October and January-February.

24. In your room there are books on the right hand side as you enter.

25. The room has recently been changed around.

26. There is a book or notebook in which you have been writing. (If not, skip to Item 33.)

27. You have had misgivings as to being able to make something fine of this writing.

28. It is unfinished.

29. The writing has been put away. (If not, skip to Item 31.)

30. In a drawer.

31. This writing is almost unknown except to yourself.

32. You seldom if ever speak about it.

only it is not loose—they are bound. This writing may be quite unknown to anyone at all and it may not be spoken of to anyone at all. It is not on a table or not in full view, but as though put away in a drawer. There had been some setback with that writing but I do feel that the continuity of that is quite important and that it might be finished.

There is also a little—it has no bearing whatever on the writing—it is like a little dog. It is grotesque—a toy—and given by a lady around Christmas season and is sitting up. It is a toy of sentiment. It is a little grotesque. It seems as though it is resting on papers. It would seem to me to be rather like a room in which one might study as well as sleep. I hope I have made that clear.

I would like to tell you a little of the peoples who have passed. I have a very definite feeling that the communicant will do very much in this subject of—how you call it—parapsychology. There is no doubt at all there are certain leanings already, and the subject is not participated in only but has a very deep interest. I am finding somebody conscientious in regard for truth, one who has quite a little knowledge along this particular line as well as a very large interest.

Of the people who have passed

33. There is a toy dog in your room. (If not, skip to Item 40.)
34. It is grotesque.
35. Given to you. (If not, skip to Item 38.)
36. By a lady.
37. You received the gift around Christmas of this past year.
38. The dog is sitting up.
39. It is resting on papers.
40. You both study and sleep in your room.

41. You have a brother de-

over I want to tell you of a brother who must have gone at an early age. I think perhaps very soon after babyhood. The only thing in connection is the name like Robert or Rupert, and must have perhaps been only a memory to the communicant.

I have, however, two people who have passed quite recently. There is somebody of the name Greely, I think, who has been connected with education and a very good friend of the family. The name Walter comes to me in connection with that one and he may have passed out at the age of 72, and in the past two years. He is not a relative but a close friend of the family.

There is a lady who went at about the age of 46 or 47. She was very round of figure, jolly, clear skin, gray eyes, and dark hair. She was extremely feminine in all qualities. She loved flowers and had a great love for beauty. To know this woman was to have immediately liked her. She had a great sense of humor, laughter and joy. She seems to be connected with the communicant on the maternal side of the family. The name Marjorie comes to me and evidently a sister Beth who is living.

- ceased. (If not, skip to Item 45.)
42. He died very young.
  43. His name was Robert or Rupert.
  44. You recall him faintly—he “must perhaps have been only a memory” to you.
  45. There is a man, deceased, who was connected with you. (If not, skip to Item 52.)
  46. His name was Greely.
  47. He was connected with Education.
  48. Not a relative but a very good friend of the family.
  49. His name was Walter.
  50. Died at the age 70-74.
  51. His death took place within the last two years.
  52. There is a lady, deceased, who was connected with you. (If not, skip to Item 67.)
  53. She died at the age 44-49.
  54. Very round of figure.
  55. Jolly.
  56. Clear-skinned.
  57. Gray eyes.
  58. Dark hair.
  59. “She was extremely feminine in all qualities.”
  60. “She loved flowers and had a great love for beauty.”
  61. “To know this woman was to have immediately liked her.”
  62. A great sense of humor.

Of the communicant's mother I feel a great deal of strength, great ability, and a woman of very great power of persuasion, a woman who possibly had an operation as though on the gall bladder or appendix not so long ago. There was a good deal of trouble and an operation had taken place. She has been very much better. She has had trouble about getting a little more flesh on, but that will adjust itself and it is very definite that she is in much better health now than she had been for some considerable time.

I have a feeling that there is a friend of the mother's not long passed over of the name Glenn as well as I can hear—it may be Glenister. I hear only Glenn.

I want to tell you that there is also a feeling that there must have been little personalities in this family, and there is one who they have revered very much and one who was also a great law-giver. He gives to me the impression of having English blood and I think he had a good deal to do with—how you call—state law. This

63. Connected with you on the maternal side of your family.
64. Her name was Marjorie.
65. She has a living sister. (If not, skip to Item 67.)
66. The sister's name is Beth.
67. Your mother is living. (If not, skip to Item 74.)
68. She is a woman of great strength and ability.
69. She has great persuasive power.
70. She underwent a surgical operation recently.
71. She has had a great deal of sickness.
72. Her health is much better at present.
73. She has been trying to gain weight.
74. A friend of your mother's is deceased. (If not, skip to Item 77.)
75. He or she has not long been dead.
76. His or her name was Glenn or Glenister.
77. There is a man deceased who was connected with your family. (If not, skip to Item 101.)
78. He was much revered.
79. A great law-giver.
80. Of English blood.
81. He had to do with state law.
82. He is often spoken about in the family.

man is often spoken about in the family, and I think the communicant will not have much difficulty in recognizing a picture I see of him at any rate. The picture shows him wearing the side whiskers, not the beard but the side whiskers. He had gray eyes, a particularly fine forehead and bushy brows. I hear the name Austerly or Austerly in connection with him and also the name George. This man would seem to me to have given a great incentive to the members of the family. This George may be a grandfather but at any rate he is on the paternal side. I hear the names of Herbert, Arthur, John and James. James is his own brother, and Herbert, Arthur and John must be sons. I would think that Herbert and Arthur might live, but the John no. I would imagine that behind this family there may be quite a little German blood, but there is also a distinct English strain as well. So I find besides the legal there were also soldiers.

And I have a feeling that in the house of the communicant—not in the room I spoke of, please, there are pictures on the paternal side of both German and English forebears. When I speak of pic-

83. There is a picture of him which you know well. (If not, skip to Item 85.)
84. In it he is wearing side whiskers.
85. He had gray eyes.
86. A particularly fine forehead.
87. Bushy brows.
88. One of his names was Austerly.
89. One of his names was George.
90. He gave great incentive to members of the family.
91. He was your grandfather.
92. Related to you on the paternal side.
93. He has a deceased brother. (If not, skip to Item 95.)
94. The brother's name is James.
95. He has a son, living. (If not, skip to Item 97.)
96. The son's name is Herbert.
97. He has a son living. (If not, skip to Item 99.)
98. The son's name is Arthur.
99. He has a son, deceased. (If not, skip to Item 101.)
100. Named John.
101. There is both German and English blood in your family.
102. Some of your relatives, living or deceased, were soldiers.
103. In your house (home)—not in the room referred to in Item 1—there are pictures of both German and English ancestors. (If not, skip to Item 106.)

tures I should make myself clear and say paintings.

There is also a sword in a case in this house connected with somebody of the name Stuart (Stewart).

I would feel that I would speak also of the gentleman of the photograph whom I spoke of before and ask you to make note of the fact I think at one time he must have devoted a great deal of time to fishing. There are pictures suggestive of this which I see in the library of the communicant's house. Now "library" may not be the correct term but a three-quarters library. That is the way it looked to me—as though that was its destiny.

In speaking of the communicant I can only speak of a lady who has had a great deal to do with the fortune of the family. She was a woman of great character, domination and beauty. There is a picture of her amongst some of the pictures—a picture of her taken young. She must have had a sister though the sister is not shown to me at all in any of these pictures. This lady I think to be the grandmother was very definitely a great

104. These people are related to you on the paternal side of your family.

105. The pictures—some, at least—are painted portraits.

106. There is a sword in your house (home). (If not, skip to Item 109.)

107. The sword is kept in a case.

108. It is connected with someone of the name Stuart (Stewart).

109. The gentleman of the photograph, Item 77, once did much fishing. (If not, skip to Item 111.)

110. There are pictures in the library of your house which suggest his fishing days.

111. Your grandmother who is deceased most nearly fits the description below.

112. She was "very definitely a very patriarchal force."

113. She died at the age 79-84.

114. Her death occurred within the past five years.

115. Her home was in the north.

116. At one time she was connected with Cambridge.

117. She is your maternal grandmother.

patriarchal force. I think she must have been about the age of 81 or 82 when she passed and I do not think it has been more than 4 or 5 years since her passing. I have the sense of a lady who had lived rather in the north and not in the south. She was connected with Cambridge at one time very definitely. She is, I would imagine, maternally connected. She seems to have had a great love for a number of children and there are times when in a very large garden I see five children. These five children must be very closely connected with the communicant's early life.

This lady had a very large garden. She had a particular aversion to animals and a great love for birds. The communicant has a ring of hers that seems to have an exceedingly fine pearl. This lady I must say has a strong spiritual influence over the communicant. Now the communicant will not agree with me because the communicant's attitude is impartial. Whatever difficulty the communicant has that simply reaches him alone is always helped in some way that seems miraculous.

The communicant has a dearly beloved sister and I feel also may have two brothers. I have no sense of them being passed away. I think they are here.

That is all I can tell you actively of the people I see except the

118. She had a great love for a number of children.

119. Five children in a very large garden draws a picture which is significant for one period of her life. (If not, skip to Item 121.)

120. These five children were closely connected with your early life.

121. The lady of Item 111 had a very large garden.

122. She had a strong aversion for walking animals.

123. She had a strong liking for birds.

124. You have a ring which belonged to her. (If not, skip to Item 126.)

125. It has a fine pearl.

126. One of the lady's names is Martha.



name of the older lady is suggested to me and I am hearing the name Martha Jane. There are many stories of hearing a Granny being one of the central characters though left.

The communicant is definitely at the later stage of life, likely to go to Europe. There is travel, study and much work. The communicant has great fundamental drive, and that drive may of a necessity make many more changes than the communicant can be aware of at this moment. There is a definite attitude toward Europe.

I would ask the communicant to do just one little thing and that is to think of a friend of the name of Young connected with a gentleman who may be studying. Quite closely of the name of Young there is one who comes to us and whom you may in your notation speak of as the intruder. I would feel this Young was a clergyman, dying about six or seven years ago. And the name Charles and Sheppard are given in connection with him. The communicant will have some friend who will probably recognize this one quite definitely.

There are many eulogies I could continue to heap upon this communicant as to the ability to make good in a very quiet and beautiful manner, but I think I have said enough and I do not see any

127. One of her names is Jane.  
128. Her names in order are Martha Jane.

129. You have a strong desire to go to Europe.

130. There is a man, deceased who is known to you but who is more closely connected with a gentleman friend of yours who is engaged at present in study. (If not, skip to Item 136.)

131. This deceased man's name is Young.

132. He was a clergyman.

133. Died 6 or 7 years ago.

134. One of his names is Charles.

135. One of his names is Sheppard.

more personalities who have passed on, nor do I get very many more in way of contact.

The communicant has recently been reading a book of Wells. It is like unto biology. On page 33 you will find in the last two paragraphs an amusing reference to a troublesome situation akin as it were to the communicant's little worried area. That is all I can say to you. You know, as I spoke to you at first of the communicant's being a little worried over certain plans, I always have the sense of the communicant feeling "I don't know if I can really carry on." These paragraphs are a parallel of just such a situation. I hope these things will have a particular bearing. I thank you and that is all. (Thank you, Uvani.)

I do definitely want to speak of a sister. I would feel that the sister is very, very closely related to the communicant, and I would think that the sister I speak of is not at all happy. Indeed, the communicant would seem to, if not be in possession of news at the moment, will be in possession of a letter from that sister speaking of this unhappiness. The letter would seem to come from the East, and I have a very definite feeling that it may speak of marital unhappiness.

136. You have recently been reading a book by Wells. (If not, skip to Item 139.)

137. A book related to biology.

138. On page 33 of this book, the last 2 paragraphs contain a reference to a troublesome situation akin to a worrying condition of your own.

139. You have a sister living. (If not, skip the remaining items.)

140. She is very, very closely related to you.

141. She is not at all happy.

142. You have had a letter from her speaking of this unhappiness.

143. It spoke of marital unhappiness.

[APPENDIX B OMITTED]

## APPENDIX C

These instructions were given to subjects in Series II when they came for their experiments with Mrs. Garrett:

### INSTRUCTIONS TO SUBJECTS

In conducting these experiments, for which you are one of the subjects, precautions are being taken to achieve two ends. One aim is to prevent the sensitive's being able to use knowledge secured through sensory channels or by rational inference. The other aim is to secure material which the subjects may be asked to evaluate for applicability to themselves without having to guard against personal bias. In order to help achieve these ends, you are asked please to follow as closely as you can the following suggestions:

Select and bring to the experiment *one* personal object which belonged to or was connected with someone deceased who was closely associated with you. In coming for your appointment, come to the fourth floor of the Medical Building and turn to the right. Continue around the hallway to the right until you come to Room M445, where the secretary will meet you. The secretary will direct you into a room adjoining that in which the sensitive will be in trance. You may talk freely with the secretary while in Room M445, *but please do not speak at all from the time of leaving this room until you return to it.* Your not speaking will control the risk of your being identified by the sensitive, either individually or as to age and sex.

You will sit alone in the room to which you are conducted for the experiment. The fan will be running to drown out any words of the sensitive which might otherwise serve to help you identify later the material of your experiment. Please remain in the chair to which you are shown and do not try to understand what is being said in the adjoining room. Your indulgence in the matter of having to ask you to wait until later for knowledge of your experiment must be asked for the sake of safe evaluative method.

During the time of the experiment, try to assume an attitude as favorable as possible to the spirit hypothesis. Sit and think of deceased ones who were close to you and from whom you would like—if possible—to have information of their present existence. Try to put skepticism aside for the time and enter into the spirit of the

experiment. This sympathetic, expectant attitude may produce better results, if only because it places you in a position comparable to that of the telepathic agent. The aim of the present series of experiments is to attempt to evaluate trance material for extra-sensory knowledge; no attempt is made to rule out here telepathy and clairvoyance as possible explanations, alongside the spirit hypothesis, of this knowledge if its presence is demonstrated.

Immediately after the experiment, you will be conducted back into Room M445 and will be asked to go away at once by the route followed in coming in. This precaution is taken in order that the sensitive may not see who the subjects have been, as otherwise she might make guesses as to who later ones will be by inference from the list of those people whom she may know to be generally eligible to serve in this capacity.

#### APPENDIX D

The following statement was mailed to all subjects in Series II with the first analyzed record which they were asked to apply to themselves.

##### INTRODUCTORY STATEMENT

In connection with the recent experiment with Mrs. Eileen Garrett, you are asked to read this statement concerning the procedure to be followed in checking upon the material secured.

This study is in most respects a repetition of the one which was undertaken last year. The previous study gave results that were "statistically significant" of other than chance factors. The question left at issue is whether these factors are reducible to some kind of extra-sensory perception on the part of the sensitive. That is, to some mode of "supernormality," as it is called. If the evidence should, with rigidly controlled experimentation and strictly scientific evaluation, favor the extra-sensory perception hypothesis, it will then be in order in later research to choose, if possible, among the various specific hypotheses consistent with this general one.

The important additional control which was introduced in this later experiment was that the subjects were not permitted to hear their material presented by the sensitive. This condition controls also the factor of possible individual differences in attitude in scoring dif-

ferent material by the subjects. This end was the main consideration in preventing the subjects from hearing the sensitive. As desirable and necessary as this end is, it seems that never before has any adequate control of the personal factor been attempted. To further insure the highest obtainable degree of uniformity of attitude throughout, you are requested to score each body of material as though it contained the results of your own single "sitting." The procedure being followed is to assign arbitrary reference symbols to each experiment and to submit the material of each experiment as a unit for evaluation as to applicability to each subject's circumstances. References to distinct personalities occurring within the same experiment or "sitting" can not be checked for applicability to one person but must be taken to indicate separate persons; however, references in different experiments may be applied to the same person as often as they seem best to fit that one.

Beginning immediately, the itemized results of the different experiments will be sent in lots of one to three, at one- or two- day intervals, to all the subjects who participated in the series. The importance of having each subject score each body of material promptly (if possible on the day received or not later than the day thereafter) and return it at once can not be overemphasized. For a subject to allow himself to depart from this rule will mean inevitable accumulation of bulky material on receipt of other installments. Needless to say, no subject can be told which is his experiment until all subjects have returned the last scored set of their results. This, again, will be appreciated as a necessary step to take for the sake of scientific precautions. Please remember, therefore, that if you allow yourself to delay in attending promptly to the scoring of all the material for applicability to yourself you may be preventing others who are more eager than yourself from knowing which is their own "sitting" and what the value of it is. This information will be sent out (in each case, of course, only to the subject actually concerned) as soon as all participants have returned all sets scored for applicability to themselves.

Please observe the following rules in scoring:

1. Answer strictly any item which contains a specific reference to age (to the last birthday), sex, and whether a person is living or dead.

2. Use the following symbols for scoring: ( $\checkmark$ ) a check to indicate *yes*; ( $\times$ ) a cross to indicate *no*; (?) a question mark to indicate that you do not know the answer.
3. If you can make inquiries to advantage, do so, and place a circle around the answers given by others.
4. Descriptions of purported communicating personalities referred to as "related to" are to be checked, in each case, for those relatives for whom the descriptions are most appropriate. Descriptions of persons referred to as "connected with" are to be checked for those persons, whether related to you or not, who seem most clearly to be indicated.
5. As you finish each lot of experiments, return them at your earliest convenience.
6. Read each section (indicated by Roman numerals) before marking.

#### APPENDIX E

*Tables of Values for Substitution in the Saltmarsh-Soal Formulas.\**

p	Correct $-(\log p \cdot q)$	Variance $(p \cdot q \cdot [\log p]^2)$	Incorrect $+(p \cdot \log p)$
.00	-----	-----	-----
.01	1.9800	.0396	-.0200
.02	1.6650	.0566	-.0340
.03	1.4772	.0673	-.0457
.04	1.3420	.0750	-.0559

\* The use of this table may be explained as follows. The first column (p) contains the ratios of general probability from .00 (items true for nobody in general) to 1.00 (items true for everybody in general). If an item is true for the actual subject, the value in the "correct" column is used; if false for the actual subject, the negative value of the "incorrect" column is used. These values for all items are added algebraically to give the score of a record. For each item also, whether true or false, the figure in the variance column is used. The square root of the sum of these entries gives the standard deviation (SD). The SD is then divided into the score to give the critical ratio, and this may be stated, by the use of a probability table, in terms of the odds against the result's being one of mere chance. [For the statistically-minded reader, the mathematical expressions under the column headings may be sufficient to explain the derivation and use of the figures. Those wishing to see a full statement of the evaluative method should consult the original article in *Proceedings, S.P.R.*, 1929-1931, vol. 39, pp. 266-271.]

p	Correct $-(\log p \cdot q)$	Variance $(p \cdot q \cdot [\log p]^2)$	Incorrect $+(p \cdot \log p)$
.05	1.2360	.0804	-.0651
.06	1.1485	.0842	-.0733
.07	1.0741	.0868	-.0808
.08	1.0091	.0886	-.0878
.09	0.9517	.0896	-.0941
.10	0.9000	.0900	-.1000
.11	0.8531	.0900	-.1054
.12	0.8103	.0895	-.1104
.13	0.7709	.0888	-.1152
.14	0.7344	.0878	-.1195
.15	0.7003	.0865	-.1236
.16	0.6686	.0851	-.1273
.17	0.6388	.0836	-.1308
.18	0.6107	.0819	-.1340
.19	0.5842	.0800	-.1370
.20	0.5592	.0782	-.1398
.21	0.5355	.0762	-.1423
.22	0.5129	.0742	-.1447
.23	0.4915	.0722	-.1468
.24	0.4710	.0701	-.1488
.25	0.4516	.0680	-.1505
.26	0.4329	.0658	-.1521
.27	0.4151	.0637	-.1535
.28	0.3980	.0616	-.1548
.29	0.3817	.0595	-.1559
.30	0.3660	.0574	-.1569
.31	0.3509	.0553	-.1577
.32	0.3365	.0533	-.1584
.33	0.3226	.0513	-.1589
.34	0.3092	.0493	-.1593
.35	0.2963	.0473	-.1596
.36	0.2840	.0454	-.1597
.37	0.2720	.0435	-.1598
.38	0.2605	.0416	-.1597
.39	0.2494	.0398	-.1595
.40	0.2387	.0380	-.1592

P	Correct $-(\log p \cdot q)$	Variance $(p \cdot q \cdot [\log p]^2)$	Incorrect $+(p \cdot \log p)$
.41	0.2284	.0363	-.1588
.42	0.2185	.0346	-.1583
.43	0.2089	.0329	-.1576
.44	0.1996	.0313	-.1569
.45	0.1907	.0298	-.1561
.46	0.1821	.0282	-.1551
.47	0.1738	.0268	-.1541
.48	0.1658	.0254	-.1530
.49	0.1580	.0240	-.1518
.50	0.1505	.0227	-.1505
.51	0.1434	.0214	-.1491
.52	0.1363	.0201	-.1477
.53	0.1296	.0189	-.1461
.54	0.1231	.0179	-.1445
.55	0.1168	.0167	-.1428
.56	0.1108	.0155	-.1410
.57	0.1050	.0146	-.1391
.58	0.0994	.0136	-.1372
.59	0.0941	.0127	-.1353
.60	0.0887	.0118	-.1330
.61	0.0837	.0110	-.1310
.62	0.0789	.0102	-.1287
.63	0.0743	.0094	-.1264
.64	0.0698	.0087	-.1240
.65	0.0655	.0080	-.1216
.66	0.0614	.0073	-.1191
.67	0.0574	.0067	-.1165
.68	0.0536	.0061	-.1139
.69	0.0500	.0056	-.1112
.70	0.0465	.0050	-.1084
.71	0.0431	.0046	-.1056
.72	0.0400	.0041	-.1027
.73	0.0369	.0037	-.0998
.74	0.0340	.0033	-.0969
.75	0.0312	.0029	-.0937
.76	0.0286	.0026	-.0906



p	Correct $-(\log p \cdot q)$	Variance $(p \cdot q \cdot [\log p]^2)$	Incorrect $+(p \cdot \log p)$
.77	0.0261	.0023	-.0874
.78	0.0237	.0020	-.0842
.79	0.0215	.0017	-.0809
.80	0.0194	.0015	-.0775
.81	0.0174	.0013	-.0741
.82	0.0155	.0011	-.0707
.83	0.0138	.0009	-.0680
.84	0.0121	.0008	-.0636
.85	0.0106	.0006	-.0600
.86	0.0092	.0005	-.0563
.87	0.0079	.0004	-.0526
.88	0.0067	.0003	-.0488
.89	0.0056	.0003	-.0450
.90	0.0046	.0002	-.0412
.91	0.0037	.0001	-.0373
.92	0.0029	.0001	-.0333
.93	0.0022	.0001	-.0293
.94	0.0016	.0000	-.0253
.95	0.0011	.0000	-.0212
.96	0.0007	.0000	-.0170
.97	0.0004	.0000	-.0128
.98	0.0002	.0000	-.0086
.99	0.0000	.0000	-.0044
1.00	0.0000	.0000	-.0000

### III

## APPRAISING VERBAL TEST MATERIAL IN PARAPSYCHOLOGY<sup>1</sup>

### INTRODUCTION

This will chiefly be an account of steps that have been taken in the Parapsychology Laboratory [of Duke University] over a period of fifteen years toward the development of a method of evaluating verbal material obtained in tests of psi abilities. It will be a report on a search for a method, a search that has resulted in definite progress even though it is still not complete. Thus far in working with responses given in the form of descriptive statements, the investigators have been trying mainly to find a way of appraising the results, and relatively little attention has been paid to obtaining the best possible subjects for such psi tests and to providing the most favorable psychological conditions.

Verbal material is one of the forms in which ESP has been most frequently reported to occur. If an adequate method can be found, a careful study of this type of response obtained under suitable test conditions would obviously have wide application and great value. The type of word material that has received the most attention in parapsychology has been the utterances of "mediums." These statements have been studied chiefly from the point of view of their bearing upon the hypothesis of spirit agency. The study of mediumship depends largely on a method for accurately evaluating verbal material, and we may say that progress in such study has long awaited the development of methods such as we are seeking.

---

<sup>1</sup> This section presents the study published in the *Journal of Parapsychology*, vol. 12, December, 1948, by J. G. Pratt and W. R. Birge (deceased). The part of the original report that traces earlier stages of methods for dealing with verbal material in parapsychology is omitted since this has been adequately covered in the preceding section. The Pratt-Birge study is reprinted with the permission of the Johnson Reprint Company.

## METHODS TRIED MORE RECENTLY

Since the publication of the preliminary report on the assessment of mediumistic utterances in 1936, the Parapsychology Laboratory has continued the efforts to improve the methods for dealing with verbal test material. During this period we have worked almost exclusively with the results obtained in token object tests. We used a number of subjects who thought they might be able to demonstrate their psi abilities in this way. The objects were always securely wrapped and sealed to avoid the possibility of inferring anything from them. In most instances, the cooperators submitted objects that were keepsakes of high sentimental value because of their connection with someone deceased. In this way we were able to work more effectively with subjects who in their experiences along these lines had been oriented toward the spiritistic hypothesis.

As we stated earlier, the emphasis thus far has been kept on the development of a suitable way of appraising verbal test material, and we shall continue to limit ourselves in this paper to this question. The results, which have not been significant as a whole, will be presented only as needed to illustrate a method. Nor shall we be concerned at this time with the details of procedure in these token object tests that do not directly affect the assessment technique. Suffice it to say that precautions have been taken to insure that the subjects should not obtain any clues regarding the identity of the cooperators and the original owners of the token objects.

In addition to the shortcomings discussed earlier, all the methods already described suffered from the practical difficulty of being cumbersome to apply. This difficulty became especially great in the methods that were first tried in the Parapsychology Laboratory. The use of control checking of the material by the cooperators led to the necessity of treating a larger number of test records at a time in order to increase the reliability of the item probability ratios. As work along these lines was resumed, it was with a clearer recognition of the fact that the method of assessment would have to be one that investigators would find practicable.

Another practical difficulty was that some of the methods had been wasteful of the subject's utterances themselves. A literal interpretation of the statements had been followed, and as a consequence

many of the items were excluded from the final evaluation as having no relevance for the cooperator for whom they were intended. [In the study presented in the preceding section, the word *subject* referred to the sitter or the person for whom the utterances were intended (here called *cooperator*), whereas *subject* in the present study refers to the sensitive or medium, the person whose abilities are under investigation.—Editor.]

In 1944 C. E. Stuart proposed a method that seemed to overcome these two practical difficulties to some degree. One suggestion he made was that when the records were broken down into items the wording used should be such that each cooperator would be able to mark a larger percentage of the statements as either correct or incorrect for himself. Stuart took the view that the subject might be giving correct information in his utterances, but that he might be directing his remarks less accurately than a literal interpretation would suggest. Such things as mistaken identity or the mixing up of two or more individuals connected with the cooperator might be occurring. Thus, suppose the subject said of a particular cooperator that he had a deceased uncle and that this uncle had been fond of smoking a corncob pipe. According to the earlier methods of itemizing the material, the descriptive item about the corncob pipe would necessarily be considered of no relevance to the cooperator if he did not have a deceased uncle. With Stuart's method, the item might still be considered applicable provided the cooperator had a deceased father or some other relative who fitted the description. For example, this particular item might be presented in two parts, in some such manner as follows: "You have a deceased relative who used to be fond of smoking a corncob pipe ( )." "This relative was your uncle ( )." This broader basis for the interpretation of the material proposed by Stuart was in keeping with the view that the subject's verbal responses might be even more "free" than the actual words of the subject seemed to imply.

Stuart also suggested a method of evaluation that required fewer cooperators for an experiment. He proposed five token object tests as the standard number. An ingenious evaluative procedure was devised that still made use of empirical ratios but did so with only five cooperators. As in the earlier work, each cooperator was asked to score all the records without knowing which one was his own. Stuart

selected for evaluation only those items to which four cooperators had responded with "no" and one had responded with "yes." On a purely chance basis, each of these items has a one-fifth probability that the person who scored the item "yes" would be the one for whom it was intended. If the cooperator who scored the item "yes" was the same cooperator for whom the subject intended that item, it was counted a hit. Otherwise it was a miss. The statistical significance of a particular record as well as of a series of five records could then be computed in exactly the same manner as that in which the deviation for a particular number of trials in a standard ESP card test is measured.

Stuart had not made a final formulation of his method before his untimely death in March of 1947. His procedure as described here offered definite advantages of the sort that he was trying to achieve. There were still some shortcomings in the procedure, however, as Stuart himself was well aware. One of these was the fact that the method, in spite of achieving a much higher percentage of judgments of the material from the cooperators, still allowed only a small amount of the information given by the subject to enter into the final calculation of the results. Another objection, one of an even more fundamental character, was that the statistical evaluation used still involved the assumption that the items were independent of one another. This requirement concerning item independence was not met in this procedure any more than it had been in the methods already described.

Stuart also made an effort to apply to verbal material the evaluative technique which he had developed for free response tests based upon pictures as targets. This is the preferential matching method that is now widely used in drawing tests. There seemed to be no a priori reason why this method should not be used for free verbal material.

In applying this method, Stuart broke down each record in a set of five token object readings into separate paragraphs dealing with distinct topics or personalities. The records were then sent to the cooperators in units of five paragraphs, including one paragraph from each of the original five records. The cooperators were asked to rank each paragraph in a unit according to its applicability to their circumstances. This method was later discarded as too wasteful,

since it lumped together all the information in a paragraph as a single trial by the subject.

After Stuart's death, one of the writers (W.R.B.) for a time assumed an active role in the effort to improve the methods for handling free verbal material. Still another new approach was tried. As before, the information in a set of five records was itemized and submitted to the five cooperators without their knowing which were their own records. For the evaluation of the results, only those items in all the material that were checked by each cooperator were counted as "trials" made by him. Of these trials, those check marks which came within his own record were considered as hits. If the records were of the same length, in the sense that each one offered the same number of items to be accepted or rejected, the expectation was that the number of check marks made by a cooperator within his *own* record would be one-fifth of the total number of his check marks in *all five* records. Any observed tendency for a cooperator to check a larger number of items correct within his own record could be evaluated in terms of the deviation from the expected number by the same formula as Stuart had used—the one that is commonly applied in ESP card work.

This method did seem to represent a definite step forward. It was a more simple and direct approach to the problem than the procedure Stuart had proposed, and it also enabled the investigator to make use of all the items that the cooperators checked as correct for themselves. Its shortcomings were that: (1) like the earlier procedures, it assumed an independence of items; (2) records of different lengths presented a special difficulty in that they could not be used without changing the  $1/5$  probability basis of evaluation required for the statistical analysis. The practice proposed was that of itemizing all the records and then dropping off items at the end so that all the records would have the same number of items as the shortest one. This device again involved the wasteful necessity of leaving out some of the material.

### THE CURRENT PROCEDURE

Up until this time all of the methods that had been tried, except that of treating paragraphs as a whole by means of the preferential matching method, involved editing of the free verbal material, or

breaking it up into items before it was scored by the cooperators. The practice in most of the methods of assessment tried had been to single out the items and to do some paraphrasing of the subject's utterances in order to help the cooperators in their marking of the material. The idea occurred to one of us (J.G.P.) that keeping the exact words used by the subject would offer several distinct advantages. In the first place, this procedure would save a great deal of time. Furthermore, it would avoid the danger of misinterpreting the verbal material in a manner that might affect the scoring. This danger, of course, becomes serious only in case the person who itemizes the material is acquainted with the cooperators. It seemed that it might be best to avoid, if possible, all "editing" of the records. Experience with these records suggested that it might be possible to achieve all the advantages of itemization simply by inserting parentheses ( ) wherever the subject made a remark that introduced a new thought or qualified a statement in any way. The only judgment required was in deciding where to insert one of these checking points in the material.

All of the foregoing methods for evaluating verbal test material in terms of separate items assumed an independence among the items that may not have existed. For the items within a record to be truly independent, the fact that a cooperator has checked a particular item as correct or incorrect should have no relationship to the manner in which he checks the remaining items in the record. There are two main reasons why the items cannot be considered independent in this sense. First of all, descriptive accounts tend toward some degree of self-consistency: items pertaining to the same topic are likely to be highly related. It follows that, if a particular item is correct or incorrect for a cooperator, an indeterminate number of other items are likely to be checked as correct or incorrect purely as a consequence of this interrelationship or self-consistency.

The second factor undermining the assumption of independence of items is the fact that, quite apart from the self-consistency of a record, the answers made by the cooperators may not be independent of one another. Thus a cooperator may form the general opinion that a record is or is not meant for him on the basis of a few introductory items. Through a perseverative tendency, the cooperator's

scoring of the remaining items in the record may show a general effect of this opinion.

The method currently being tried out in the Parapsychology Laboratory does not assume independence of items. The logic of this new method may be explained by an illustration drawn from familiar areas of ESP research. Investigators have long been aware that data obtained from group tests may not be interpretable in terms of the same statistical analyses as apply to the data of individual tests. In group tests, a number of subjects attempt to call the same targets, and there is no way to be sure that any nonrandomness found in the calls made by different members of the group could be attributed only to the special ability (ESP) being tested. For example, subjects in a group might exhibit a tendency to start their calls with a certain symbol, or they might show the same symbol preference in their series of calls because of cultural or environmental influences or for some other reason having nothing to do with ESP ability. Any such group pattern or characteristic of response would not change the expected average number of hits, but it would affect the variance, the distribution of total scores about the mean for the group. The data could not properly be evaluated on the binomial hypothesis, which assumes that the calls are independent. For the variance of the binomial hypothesis to apply, it is necessary that at least one of the series of events to be compared—either the cards or the calls—be random. This condition holds in a test in which a single subject calls a particular card order only once; for the test to continue, a new card order is provided. In this case the use of a new order of target symbols for each run meets the requirement for randomness. The difficulty in the group situation is that the same card order is used over and over again in checking the calls of the individual subjects.

What statistical procedure might be used to evaluate group tests without making the assumption of independence of results when many call sequences are compared with the same card order? Greville worked out and published a solution to this problem.<sup>2</sup> His procedure involves taking the actual distribution of calls in the separate trial positions (the calls actually made on each target symbol) as the given data. For example, consider the case of 100 subjects attempting

---

<sup>2</sup> Greville, T. N. E., "On multiple matching with one variable deck." *Annals of Mathematical Statistics*, 1944, vol. 15, pp. 432-434.



to call through a single random order of 25 ESP symbols. The results would be evaluated in terms of 25 trials, not as 2,500, the actual number of calls made. The statistical question asked is as follows: Given the particular distribution of symbols in the calls as observed in each of the 25 trial positions, what is the probability that this fixed distribution would give a total score as high as the one made by the group? The total score on the test would be found by comparing the distribution of calls with the order of symbols in the target deck, this order being a random arrangement which is only one of a large number of possible permutations of the 25 symbols. For each of these permutations of the target deck, a score could be obtained by checking the order of symbols against the "fixed" or observed distribution of the subjects' calls. If the deck were arranged in every possible way and the scores were checked for every permutation, it would be possible in theory, though impracticable in fact, to work out the mean score and the variance of the distribution of all these scores. The score made by the group in the actual test (that is, the hits obtained when the observed distribution of calls is checked against the particular random order of 25 card symbols used) could then be measured in terms of its deviation from the mean and the standard deviation of the entire distribution.

Greville developed the formulas for applying this particular test of significance without the necessity of actually permuting the target order through all its possible arrangements. He considered two general situations: first, that in which the target order represents a truly random selection from among the choice possibilities offered; secondly, that in which a closed deck is used, presenting an equal number of all the possibilities with a random arrangement such as might be given by adequate shuffling. Only the closed deck situation is relevant to the problem of evaluating verbal material.

In thinking over the difficulty which the interdependence of items in long descriptive accounts seemed to cause for statistical assessment, it occurred to one of the writers (J.G.P.) that the Greville method for evaluating the data of group card tests could also be applied to the results obtained from token object tests. As a means of making this application easier to follow, we shall first give another illustration of the use of the method in evaluating the results of a group test involving card calling. In this instance we shall set up

a hypothetical situation that is more closely parallel to that represented by token object tests.

Assume that a group of 100 subjects is being tested for ESP ability. They are told that the test will consist of five trials. For targets, cards bearing the common surnames Jones, Brown, Smith, Hill, and Greene will be used, each target being used only one time during the five trials. The target order is thus one of the 120 possible permutations of the five target names.

Assume, further, that the instructions are given that each subject is to make a response on any particular trial only if he feels confident of making a hit. This would lead to differences in the number of calls from trial to trial, a condition which presents no difficulty in the use of the Greville method of analysis. The method, in other words, makes possible the computation of the mean and the variance of the scores that make up the general distribution to which the actual score belongs regardless of unequal numbers of calls from trial to trial. Those who are interested in the technical statistical aspects of this method will see the reason for this from Greville's original article. Others may accept it as a statement of fact.

Assume that when the data are tabulated in terms of the frequency of calling each name on each trial, the following distribution of responses on the five trials is found:

TRIAL	SUBJECTS' RESPONSES				
	Jones	Brown	Smith	Hill	Greene
First .....	9	7	0	0	3
Second .....	3	0	0	0	12
Third .....	10	0	26	3	13
Fourth .....	2	0	0	2	12
Fifth .....	3	0	0	4	4

The Greville method takes these figures as the given data, and it thus avoids making any assumptions regarding statistical independence among these observed responses. Taking these data as they are, the Greville method enables us to find the *mean* and the *variance* of the 120 scores that would be obtained when this particular matrix of responses is checked against the 120 different orders in which the five target names might have been presented.

The reason it does not matter whether these calls were inde-

pendent or not is simply that the method makes no assumptions regarding how they came to be distributed as they are. The analysis merely assumes that the order in which the five stimulus names are presented is random, and this requirement is fully met by the conditions. For example, the names might have been those of five persons known to the group of subjects, and the wide variation in the number of names called might have reflected differences in popularity. The Greville test of significance would still give a correct probability figure showing how frequently, on a purely random basis, this highly biased set of responses on the five trials would correspond with the order of target names as well as was found to be true in the particular instance observed. If the probability figure meets the accepted criterion for statistical significance, the indication of a causal relation between the responses and the random order of target names is precisely as strong as in an experiment analyzed by any other method if the two sets of results happened to be significant at the same level.

Suppose that in this instance the random target order Brown, Hill, Smith, Greene, and Jones had been used. By summing the figures from the appropriate column for each trial, we find that the score of the group was 48. To determine whether or not this is significant, we only need to derive the mean and the variance by the Greville statistic, find the deviation of this score from the mean, and then divide the deviation by the standard deviation (square root of the variance) to arrive at a critical ratio for the test. All that this test of significance assumes is that the target order used was selected at random.

Keeping this illustration in mind, we may turn now to the question of how the Greville method of evaluation can be applied to verbal material obtained in token object tests. An example of its application to the records actually obtained in a set of five such tests will be given. For convenience, let us designate the cooperators as Jones, Brown, Smith, Hill, and Greene.

When the five records of this set were received from the subject, they were prepared for marking by the insertion of checking points within the verbatim statements. The five records were then arranged in random order and each one was given a code designation. They were then typed with sufficient copies and all five records were sent to each of the five cooperators to be checked throughout. The

cooperators were given instructions to mark the items within each record on the assumption that they were all intended for them, and to use a check mark (✓) to show a correct statement, a cross (X) to show an incorrect one, and a question mark (?) for a doubtful item. In general, the instructions were so worded as to encourage a liberal interpretation rather than a restricted one. The aim was to get the widest possible measure of application of the material to the personal circumstances of each cooperator.

After the marked records were returned by all five of the cooperators, the items checked as correct were tabulated for each one, thereby showing the number of items he marked as correct in each record. The following distribution of check marks was found:

RECORDS	COOPERATORS' CHECK MARKS				
	Jones	Brown	Smith	Hill	Greene
First .....	9	7	0	0	3
Second .....	3	0	0	0	12
Third .....	10	0	26	3	13
Fourth .....	2	0	0	2	12
Fifth .....	3	0	0	4	4

These figures mean that Jones checked nine items as correct for his circumstances in the first record, three in the second, etc. These figures, reading them horizontally, could be thought of as meaning that the first record was called "Jones" nine times, "Brown" seven times, etc. Thus the figures take on the same significance as the distribution of subjects' calls in the illustration of the group test with five cards bearing the same names as shown on page 61. In fact, we have used identical figures in both illustrations to emphasize the close similarity of the two situations from a statistical point of view.

In the case of the token object test illustrated here, the "owners" of the five records were the targets. These were presented in one of the 120 possible permutations, an order that was selected at random and kept secret from the cooperators until after they had done their checking. It is this random order of "targets" that provides the basis for applying the Greville statistic, and the evaluation gives a straightforward statement of the probability value of the observed score obtained from the way in which the cooperators distributed their check marks.

In the example we have presented, the first record was Brown's,

the second Hill's, the third Smith's, the fourth Greene's, and the fifth Jones's. The total score and the evaluation are exactly the same as for the group card test with the random order of target cards assumed in that instance.

If significant results were obtained on verbal material appraised in this manner, the interpretation would be made in the same manner as for any test of significance. The method would permit the exclusion of chance as a reasonable explanation of the results, and this is all that any statistical analysis can do. If only five records were used in an analysis, it would not be possible to obtain a P-value of less than 1/120 for a single set. When small sets are used, the results of a number of them combined might form the basis for any conclusion.

The reader who is interested in the more technical aspects of this method is referred to Greville's original article. In the appendix we have presented the basic formulas only as far as they are necessary for the evaluation of verbal material. Also, the complete evaluation of the data from the set of token object tests described in this section is presented in the appendix.

## SUGGESTIONS

The following appear to us to be the most urgent research needs:

1. It is important to find out how sensitive the Greville method is when using different numbers of records in each analysis. As some of our colleagues have suggested, sets of five tests may be too small for the best results. But for convenience and speed in handling the results, small sets are advantageous. Therefore the aim should be to keep the sets as small as possible without making the test of significance too insensitive.

2. For the purpose of further refining the statistical practices, it is important to have material that will give significant results. Ideally, this should be obtained from actual experiments with psi capacities. But if successful subjects for token object tests are not available, verbal records might be deliberately made up to fit certain cooperators in order to see what size set is to be preferred. This method has already been used with good effect in the Parapsychology Laboratory in comparing different methods of evaluation.

3. The psychological conditions, too, are important. The results

obtained thus far suggest that in this respect the tests have not provided the essential requirements for success. For example, the use of sealed token objects may be an unfavorable way of working. It might be better to use exposed token objects that would not reveal anything regarding their owners, such as similar buttons or keys. Heretofore we have been sending all the token objects for a set of tests to a subject at one time, and there have been some indications that under these conditions a confusion among the cooperators comparable to the displacement effect may have occurred on several occasions. The results suggest that it might be preferable to send out only one token object at a time. Both patience and ingenuity are needed to devise test procedures that will satisfy the statistical requirements and provide favorable psychological conditions at the same time.

### APPENDIX

The use of the Greville method in evaluating verbal material may be illustrated by the data from the set of token object tests given in the paper. We shall present the formulas in convenient computational form without going into technical questions of derivation and proof.

For the analysis, a slightly different arrangement of the figures from that shown on p. 63 will be convenient. The columns showing the number of items the cooperators said were correct for themselves may be arranged across the page in the same order as the cooperators down the left-hand margin. The marginal totals of the figures within the matrix are needed, as well as the total number of times items were marked as correct. The data as prepared for evaluation are then as follows:

COOPERATORS	COOPERATORS' CHECK MARKS					Total
	Brown	Hill	Smith	Greene	Jones	
Brown .....	7	0	0	3	9	19
Hill .....	0	0	0	12	3	15
Smith .....	0	3	26	13	10	52
Greene .....	0	2	0	12	2	16
Jones .....	0	4	0	4	3	11
Total .....	7	9	26	44	27	113

The number of items checked, or the sum of the figures in all 25 cells of the table, constitute the "calls." In this case, the number of calls is 113.

The number of "hits" is the sum of the figures on the main diagonal, or 48.

The mean number of hits expected is one-fifth of the number of calls, or 22.6.

This particular set thus gave 48 hits where 22.6 were expected, or a deviation of +25.4.

In order to compute the variance of the scores from the observed distribution, the following values are required:

The square of the number of calls ( $N^2$ ) = 12,769

The sum of the squares of the individual cells ( $\Sigma a^2$ ) = 1,439

The sum of the squares of the separate row totals ( $\Sigma r^2$ ) = 3,667

The sum of the squares of the separate column total ( $\Sigma c^2$ ) = 3,471

If  $n$  is the number of token object tests in a set, the variance is given by the expression:

$$V = \frac{1}{n^2 (n - 1)} [N^2 + n^2 (\Sigma a^2) - n (\Sigma r^2) - n (\Sigma c^2)]$$

For the present case we have:

$$\frac{1}{25 (5 - 1)} [12,769 + 25 (1,439) - 5 (3,667) - 5 (3,471)] = 130.54$$

$$\text{The SD} = \sqrt{V} = 11.43$$

$$\text{The CR of this set is therefore } \frac{+25.4}{11.34} = +2.22$$

(P = .013)

When the results from a number of sets involving the same number of cooperators are combined, the mean number of hits for the total is the sum of the means of the individual sets, and the variance is the sum of the separate variances.

#### IV

### INVESTIGATIONS BASED ON VERBAL MATERIAL: PROGRESS AND PROSPECTS

The Pratt-Birge article, presented in the preceding section, stirred up a surprising amount of interest. It was, after all, only a methodological study and incorporated new findings merely as an illustration of the evaluative procedure. Methods in science, as research *tools*, must pass the test of how they work when put to use. Yet our method received wide attention *before* it had been submitted to this test. This fact speaks for a general recognition of the importance of the problem area concerned.

But the attention given the proposal did not mean uncritical acceptance. Questions were raised about statistical aspects of the method. Did the proposed procedure escape the limitations earlier methods suffered because of interdependence among the verbal statements? Would the method exaggerate the statistical level of significance due to some artifact or some lack of refinement in the mathematical handling of the data? If the method itself was sound, what number of verbal test sessions or what number and size of groupings of sessions would be adequate for a conclusion? Some of these questions reached the stage of public discussion in the "Letters" column of the *Journal of Parapsychology*, and it seems justified to say that the judgment rendered by Dr. T. N. E. Greville largely satisfied the questioners.<sup>1</sup> He suggested that an investigation designed to cover 50 verbal records to be evaluated in this way would exceed the amount of data required for a conclusion.

The stage was set for a new era of research into mental mediumship. But no rush to do the research developed. Why not?

---

<sup>1</sup> Letter to the Editor, *Journal of Parapsychology*, 1949, vol. 13, pp. 137-138.



Quite frequently during that period, the remark was heard that the Golden Age of mediumistic research had passed. There were no longer any people available for study who had abilities comparable to those of Mrs. Piper, Mrs. Leonard, and other great mediums of the past. Perhaps this was, at the surface level, an accurate description of the situation. But a real question remains of what was cause and what was effect. Did such gifted sensitives *seem* to have become unavailable because investigators, who had turned their research attention in other directions, were no longer interested in them? Or were research workers forced to turn their attention to other problems because suitable subjects for study could no longer be found?<sup>2</sup>

A few efforts to apply the Pratt-Birge method were made, only one of which achieved significant positive results and was fully reported.<sup>3</sup> In this and one other use of the method<sup>4</sup> the investigators suggested ingenious changes in the manner of reducing the verbal material to the mathematical figures needed for final assessments. These modifications were designed to make the method more sensitive in detecting the accurate or paranormal statements in the sensitive's utterances. The suggestion was first offered by Mr. W. G. Roll (in an earlier study than the one cited) in connection with records he collected from an English sensitive.<sup>5</sup> He requested that the cooperators should give two check marks for statements that they found particularly striking in relation to their own circumstances. The investigator later counted those items as having a value of 2, while those that were "single" checked had a value of 1. Thus the value that each cooperator gave to a record was the total of all the check marks given to the items it contained, and this was the figure entered in the appropriate cell of the table for evaluation by the Greville formulas.

The results of this exploratory investigation did not reach a clear-cut level of significance, but they did suggest that this method

---

<sup>2</sup> One indication that the former was the case is the fact that Mrs. Garrett, a sensitive who was not only available but was eager to participate in scientific investigations, was used far less than her talents justified.

<sup>3</sup> Schmeidler, G. R., "Analysis and evaluation of proxy sessions with Mrs. Caroline Chapman," *Journal of Parapsychology*, 1958, vol. 22, pp. 137-155.

<sup>4</sup> Roll, W. G., "Designs for tests with free response material," *Journal of the American Society for Psychical Research*, 1962, vol. 56, pp. 184-195.

<sup>5</sup> Roll, W. G., "Theory and experiment in psychical research." B. Litt. dissertation, Oxford University, 1959.

of assigning a score of two to an outstanding item was more sensitive than the earlier method of scoring each correct item with the value of one.

Dr. Gertrude Schmeidler, in her study mentioned above, carried further this principle of differential weighting of items in the records, doing so to a marked degree. She had someone who was not acquainted with the personal circumstances of the absentee sitters (co-operators) read through the records and rate each item on a scale of one to ten for its importance or degree of uniqueness. When the checked records were returned she evaluated the results by the Pratt-Birge Method in three different ways. One was simply to count the number of items in each record that were checked correct. The second was to add up the values of the weights that had been assigned to the items that were checked correct. The third method gave special importance to those instances in which the sensitive was credited with having made two or more successive correct statements. The weights that had been assigned to these statements were multiplied together, and this product was added to the weighted values of isolated correct statements. These three methods gave three separate sets of figures for evaluation by the Greville formula, and the study was designed to see how the three would compare in statistical significance (test efficiency or sensitivity). She found that the results showed an increasing level of significance in terms of these three methods of treatment in the order in which they are described here.

The results of the studies by Roll and Schmeidler suggest that a method of taking account of differences in quality among the individual items may provide a more sensitive test of whether the utterances reflect paranormal information. After Dr. Schmeidler's report was published, the question was raised by Dr. R. H. Thouless<sup>6</sup> whether the method of using extremely large weighting values for particular items could be statistically justified. This question was put to empirical test by Dr. Greville with results that led him to conclude that the method was adequate to accommodate even this extreme degree of numerical adaptation.

Since Dr. Schmeidler's research report was published in 1958 there have been no further publications of significant results obtained

---

<sup>6</sup> In *Journal of the Society for Psychical Research*, 1958, vol. 39, p. 331.

with the Pratt-Birge method. I know of a few exploratory tests upon records obtained from sensitives without significant results, and there may be, of course, some of which I am not aware.

Perhaps this general picture of developments might appear to indicate that the method has not worked out well in practice and that it is no more than a relic for the museum of discarded parapsychological methods. Those inclined to this pessimistic viewpoint might feel further justified by the fact that the relative neglect of this quantitative method for assessing verbal material has occurred concurrently with a strong revival of interest in parapsychological evidence obtained in studies making use of verbal material. One group of investigators tried the Pratt-Birge method on their own material in the preliminary stages of their "work"<sup>7</sup> and abandoned this approach in favor of one closer to that used several decades ago by Saltmarsh. They said that their cooperators objected to scoring a large mass of material when they knew that only a limited section of it was intended for themselves. I wonder if they informed cooperators in advance that they would have to read and check a number of control items in addition to those in their own records? Certainly all reasonable measures should be tried before going back to a method that, to a serious degree, placed their findings on a debatable basis.

We cannot afford to lose sight of the proper place of statistical methods in parapsychology. In our field as in other branches of science where statistical evaluation is required, the purpose is to operate at a higher average level of accuracy in separating facts from appearances. But the statistical approach is desirable only when the material under investigation does not, by its very nature, classify itself as "fact" or "non-fact" in the absence of mathematical assessment. If there is no vagueness on each observation regarding what was taking place, the use of a statistical sieve for sorting out events is superfluous. This principle is implicitly recognized by investigators, even by those who have long leaned more heavily upon statistics.

For example, investigators who have been concerned with laboratory experiments on psychokinesis that depended upon statistical

---

<sup>7</sup> Van Asperen de Boer, S. R., Barkema, P. R., and Kappers, J., "Is it possible to induce ESP with psilocybine? An exploratory investigation," *International Journal of Neuropsychiatry*, 1966, vol. 2, pp. 447-473.

assessment of results abandon this emphasis without a backward glance when considering a poltergeist case. There are of course real questions that have to be faced in trying to decide what caused (for example) a porcelain figurine to leave a table on which it stood and shatter against a bookcase across the room, but none of the answers depend upon statistics. Far from regretting this fact, investigators welcome the freedom from the necessity of raising statistical considerations.

Are there, similarly, studies involving *verbal material* in which the research worker does not need to be concerned with statistical methods? I think that there are, and I think that recent studies of verbal material offer some good examples. But before pointing them out I should say that recognizing that they exist does not reduce the importance of having available adequate methods for the statistical evaluation of other verbal records in parapsychology where such assessment is essential. This distinction is just as valid in regard to verbal material in parapsychology as is that between the data of laboratory tests of PK and the physical effects of poltergeist phenomena.

What, then, are some examples of verbal records in parapsychology that escape the need of statistical evaluation? In citing only a few examples I do not imply that these cases are all that could qualify, nor that they are the best examples available. The instances cited are taken from the recent literature to underscore the fact that such cases form part of the current scene in parapsychology. It is reasonable to expect, therefore, that such non-statistical studies based upon verbal records will continue to interest parapsychologists in the future. Indeed, renewed interest in evidence of this kind is a relatively recent change in the field. If this marks a trend, such studies may claim a larger share of research attention over the years ahead than they do at the present time.

The first example of a non-quantitative study of verbal material is the case of Edgar Vandy.<sup>8</sup> This report deals with the verbal records obtained in a series of mediumistic sittings held in 1933 and 1934, so only as a *publication* (1957) is it a recent case. For the present purpose, it is not necessary to summarize the details of the case. It is sufficient to state that the series of sittings held with three

---

<sup>8</sup> Gay, K. (ed.), "The case of Edgar Vandy." *Journal of the Society for Psychical Research*, 1957, vol. 39, pp. 2-64.

different mediums produced a number of accurate statements of a personal, descriptive, and technical kind which left no reasonable doubt that they referred to events in the life of Edgar Vandy and to his death by drowning.

A second example from the area of mediumship is the book, *Swan on a Black Sea*, presenting a series of automatic writings by the Irish sensitive, Geraldine Cummins.<sup>9</sup> From the contents of this series of 40 scripts, it is obvious that they contain information about the English sensitive Mrs. Willett (pseudonym) who had died a short while before the automatic writings were produced. Miss Cummins was not told for whom she should attempt to do automatic writing, nor is there conclusive evidence that she knew who Mrs. Willett was in real life. On the basis of the information produced during the first six sessions she correctly surmised that the communicator was Mrs. Willett and who Mrs. Willett had been in real life. I am not suggesting that there are no questions remaining unanswered regarding the interpretation of the material presented in this book, but I am saying that questions have not been raised because of the lack of statistical appraisal of the information.

A final example of a parapsychological study in which verbal material transcends the need for mathematical treatment is the book of Dr. Ian Stevenson reporting some of his investigations of reincarnation cases.<sup>10</sup> In this volume he presents statements (including references to names, places and other facts) produced by very young children who began to claim almost as soon as they started to talk that they remembered events connected with an earlier life in another place. (Other aspects of the findings also bear upon the evidentiality of these claimed memories as applying to a specific earlier life, but for the moment we are concerned only with the accuracy of the verbal statements.) When the child's claims led to the identification of the person that he claimed he had been, and when the statements were found to coincide with the facts at many points, the similarity between statements and the previous life speaks for itself. That is to say, this is a question that does not depend upon statistics. Again, we

---

<sup>9</sup> Cummins, G., *Swan on a Black Sea*. London: Routledge and Kegan Paul, 1965.

<sup>10</sup> Stevenson, I., *Twenty Cases Suggestive of Reincarnation*. New York: American Society for Psychical Research, 1966.

are not speaking here of *interpretation*, but only of the fact that deciding whether or not a relationship of some kind exists does not require a statistical test.

Where, then, is the boundary line between verbal material that does and that which does not call for statistical evaluation? This is not a line that can be sharply defined. But it is nevertheless a border zone that clearly exists, and we can recognize basic distinctions between verbal records that clearly belong on different sides of this dividing zone. Consider the three examples cited. In each case the information is so specific and the individual statements so unique in their reference to a particular individual that they jointly could only refer to that one person.

On the other hand, often the material is of a more generally descriptive kind, and the references are to personal relationships that do not apply only to one person. Separately, each statement might apply to a large part of the general population. For such verbal material, evidentiality depends upon whether the level of accuracy of the material as regards the person for whom it is intended is consistently higher than for others. In such instances, judging the evidentiality of the statements without objective and statistical evaluation is virtually impossible. As a good illustration of material of this kind, the reader is referred back to the record of one of Mrs. Garrett's sittings reproduced on pages 34-44.

To point out that there is a difference in specificity between the information contained in that record and that found in the Vandy report or the Cummins-Willett scripts is not to cast reflections upon Mrs. Garrett's mediumship. Rather, it is only to recognize that verbal material can be arranged on a scale of specificity. The fact that much of this material is found near the "less specific" end of the scale is not to say that it has no value from the standpoint of parapsychology. Indeed, it may in some ways have more value for this reason, because vague utterances may reflect more truly the manner in which psi finds overt expression in verbal material. Indeed, in the long run we may learn more from studying information drawn from this end of the scale than from the other.

Until parapsychologists have established unquestionably their claim to a distinct field of scientific inquiry (that is, until this claim is recognized generally by scientists in other areas), we cannot afford

to be wasteful of any phenomena that appear to be relevant to our range of interests. This means that if free-ranging verbal responses are the method of choice for some sensitives for demonstrating their paranormal powers, we must be prepared to let them work this way and to find out whether they are successful.

The methodological steps already taken, as surveyed in this monograph, indicate that the need has been clearly recognized and that progress has been made toward filling this need. A method for evaluating verbal material is now available and waiting to be applied where it may be required. At the same time, this statement does not mean that the final objective in methods has been reached and that no further significant advances are to be expected. Just as research itself in science is never completed, so the methods by which research is accomplished must never be regarded as finished and frozen.<sup>11</sup>

Indeed, the publication of this monograph *does* imply that I believe the research is going to advance along lines relevant to the methods considered here during the period just ahead. I freely admit that I do so believe, and I think we can expect a reawakening of interest not only in the problem of mediumship but also in other kinds of verbal material. There will be, I feel sure, further instances in which investigators wishing to make a more intensive study of spontaneous case material from everyday life will find themselves facing the need for more clear-cut verification of the evidence. Then too, we are already witnessing a renewal of interest in studies in which the subject is asked to record his responses through unrestricted verbalization.<sup>12</sup>

Another example of such research is the program of work on telepathic dreams now in progress in the Division of Psychiatry of Maimonides Hospital.<sup>13</sup> Until now, the methods of assessment that

---

<sup>11</sup> An important new development in methodology has been published just as this monograph was going to press. This is the article by W. G. Roll and D. S. Burdick, "Statistical models for the assessment of verbal and other responses," *Journal of the American Society for Psychological Research*, 1969, vol. 63, pp. 287-302.

<sup>12</sup> Moss, T., and Gengerelli, J. A., "Telepathy and emotional stimuli: A controlled experiment," *Journal of Abnormal Psychology*, 1967, vol. 72, pp. 341-348.

<sup>13</sup> Ullman, M., Krippner, S., and Feldstein, S., "Experimentally-Induced Telepathic Dreams: Two Studies Using EEG-REM Monitory," *International Journal of Parapsychology*, 1966, vol. 7, pp. 577-603.

have been applied have not followed the lines leading to an evaluation of the results by the Greville formula. But it appears to me that there may be advantages to be gained by having the experimental records of each series of dreams checked from this point of view.

What this would involve would be simply to assign to a judge one of the target pictures without allowing him to know on which night that target was used. Then the judge would be asked simply to read through the subject's description of dreams as recorded for all of the nights and to check those items in the descriptions which could be considered accurate or relevant for his one "target" picture. Out of such a checking procedure, when carried out by people judging the material against all of the targets used, would emerge the figures that could be applied in the Greville statistical method as adapted for the evaluation of verbal material.

It remains to be seen how the level of significance obtained by this method would compare with the levels of significance obtained by judging methods used until now. The task of the judge may prove to be easier, because he would only need to keep the features of one painting and its theme in mind and to consider each statement or item in the dream descriptions as being relevant or not to his target. He would, of course, be fitting the picture not only to the dreams produced during the night when it was actually used as the target but to all the other dreams on the other nights of the series as well, and the success of the experiment would depend upon the tendency of the judges to assign more points on the real dreams than on the others that serve as controls. But these are aspects of the procedure that have been presented more fully and more clearly in the preceding section of this monograph.

As far as the Maimonides dream data are concerned, the situation seemed to be one that could be tried in practice, even by reevaluating some of the records from previous series. This would show how the tool available from the Greville method worked out in comparison with the statistical evaluation that had already been applied. The investigators kindly accepted my offer to make such a test and sent me a series of eight target pictures and the records of the dreams recorded for the eight nights on which they were used as targets, one picture a night. No dates or other markings appeared on the



pictures and typed dream records to show which ones belonged together.

The testing of this material for significance by the Pratt-Birge method was carried out by two individual scorers. One was Mr. Nils Jacobson, a Swedish medical student who was visiting the University of Virginia during the summer of 1967, and I was the other. We worked completely independently, N.J. completing the task before J.G.P. started. Each scorer took one of the eight pictures and then read through the eight dream records. Whenever a statement was found that seemed to refer to the picture a tally mark was made for that particular picture-record comparison. Then the process was repeated by choosing another picture and reading through all the records. This led in the end to an 8 x 8 matrix of 64 scores for the eight dream records compared against the eight target pictures, and each scorer's records was then calculated by the Greville method and later evaluated, after the key had been provided by the Maimonides investigators, for the targets and dreams that belonged together.

As stated above, N.J. scored the material first, and I knew before undertaking this task that he had been awarding "hits" very sparingly. I therefore decided to be generous in crediting the dreamer with making statements that fitted the picture which I was considering at the moment on the assumption that it was the target. N.J. had given a total of 221 tally marks in his picture-dream comparisons, and I ended up with a total of 444. N.J.'s data gave a somewhat better level of significance than mine. His scoring applied to the correct target-dream associations to a degree that would be expected on a chance basis only 3 times in 100, while mine was quite insignificant ( $P=.4$ ). On the other hand the method of matching by judges that the Maimonides team had been applying gave a clearly significant result ( $P = .001$ ).

This trial application of the Pratt-Birge method was thus not a success, but it is probably just as well that it was not as far as the purposes of this monograph are concerned. It serves as a warning against thinking that the use of such methods is easy, automatic, and sure-fire in outcome while providing an opportunity to recapitulate the main lessons pinpointed by the developments in methodology that have been pulled together here:

1. Since the beginning of systematic scientific study of the questions with which parapsychology is concerned, a number of the problems have been intertwined with behavior that is recorded as free-ranging verbal material.

2. In modern-day science, an investigator must choose for much of this material between ignoring it and the questions it raises or dealing with it by an objective method that clearly guards against the dangers of overinterpretation that are otherwise present.

3. There are certain basic requirements that must be met by any adequate statistical method of assessment that might be applied to such material, and the main contribution made by the two papers that have been reproduced here in Parts II and III has been to spell out and give emphasis to these needs.

4. There may be many ways (even many not yet thought of) in which verbal material in parapsychology can be evaluated while meeting all the basic statistical requirements. One way may be best for one type of material and another for a different kind of observation.

5. In some investigations it may be important to consider the experience and qualifications of the persons used to judge the verbal records. Thus the consistent success of the Maimonides dream research may depend in large part upon the judges used in the evaluation of the findings; and any unsuccessful effort at replication of those studies must take into account the possibility that the failure occurred at least in part in the evaluative phase of the research.

6. Satisfactory evaluation that yields statistically significant results does not automatically provide one final, single interpretation of the data. It only narrows the possibilities, and it is necessary for the investigators to consider how to narrow them further until (ideally) only one unambiguous explanation is left. For example, one must consider the possibility that the parapsychological aspect of the results is introduced by the act of judging the material—a sort of “projective ESP response” of those who are asked to find (for example) which target picture was “sent” by the telepathy agent during each set of dreams. Here such things as consistency (or lack of it) among judges marking the same material help to narrow the range or to tip the scales decisively toward one or the other possibility of interpretation.

In short, there is no one approved way to evaluate verbal material in parapsychology, and there is no reason to expect quick and easy answers to questions encountered in this corner of our field. Nature has not organized her phenomena for the convenience of those who would ferret out her secrets, and perhaps the mysteries connected with the unconscious working of the mind are the most baffling of all. But for many, it seems equally clear that the solutions will ultimately repay many times what it may cost in patience and perseverance to reach them.