

PROCEEDINGS

OF THE

American Society for Psychical Research

Vol. XXV

October, 1952

Studies in Extrasensory Perception Experiments Utilizing an Electronic Scoring Device

S DAVID KAHN

Contents

	Page
Acknowledgments	i
Author's note	ii
Introduction	1
List and full definition of hypotheses tested	2
Hypothesis I: Primary ESP hypothesis	4
Hypothesis II: Decline effect	12
Hypothesis III: ESP over long distances	15
Hypothesis IV: Free vs. rigid conditions	17
Hypothesis V: Attitudes	21
Subhypotheses:	
a. "Sheep" vs. "Goats"	21
b. Influence of belief that ESP actually occurs	23
c. Attitude towards personal ESP success	24
d. Confidence in ESP ability	24
e. Expansives vs. Compressives	27
f. Emotional reaction to personal ESP success	28
Hypothesis VI: Well-adjusted vs. poorly adjusted subjects	33
General summary	36
Appendix I: A Study of Decline Effects in ESP	37
S David Kahn and Ulric Neisser	
Appendix II: Example of IBM sheet used in Testing	45
Appendix III: Attitude Questionnaire	46

Published by

American Society for Psychical Research, Inc.
880 Fifth Avenue, New York 21, N. Y.

Fellowship \$25

Price \$2.00

Membership \$10



133.07
A512P
v. 25
Oct. 1952

Acknowledgments

I should like to express my gratitude to a number of individuals who gave freely of their time and knowledge in assisting me in this work. Those to whom I am especially indebted are:

Dr. Pitirim A. Sorokin, of the Harvard Department of Social Relations; Dr. Gardner Murphy of the City College of New York; Dr. Edwin B. Newman, Dr. B. F. Skinner, and Dr. F. C. Frick of the Harvard Department of Psychology, for their suggestions and criticisms on various points; Dr. J. G. Pratt, Research Associate at the Duke University Parapsychology Laboratory; Dr. Fred Mosteller, of the Harvard Department of Social Relations and Dr. T. N. E. Greville, Principal Actuarial Mathematician, National Office of Vital Statistics, for their statistical advice and assistance; Dr. J. B. Rhine for his constant encouragement and cooperation; and Mrs. E. W. Allison and Mrs. L. A. Dale of the A.S.P.R. for their invaluable editorial assistance. The writer also wishes to acknowledge gratefully the financial assistance given to the project by the Harvard Department of Psychology, and the Harvard Society of Parapsychology.

Part of the material has been presented in the *Journal of Parapsychology*, 1949, vol. 13; courteous consent to its republication in the present form is gratefully acknowledged.

Mr. Ulric Neisser of the Harvard Society of Parapsychology collaborated with the author on the actual research work, and thus shares equal responsibility for it. Other members who assisted include Mr. John Altrocchi, Mr. Herbert Fraunfelker, and Mr. Tom Plaut. The responsibility for this presentation, for the conclusions drawn from the experimental observations, and for the general point of view, lies with the author.

Author's Note

During the last twenty years, closely controlled quantitative methods have come to dominate the field of parapsychology. It has, however, been notably difficult to demonstrate extrasensory perception in the general population without first selecting the subjects on the basis of attitude or personality structure. Moreover, some critics have felt that the experimental techniques used, even in recent work, have not been fully adequate.

The following *Proceedings* represent an attempt to demonstrate the functioning of ESP in a large group of subjects by a method which may adequately invalidate previous counter-hypotheses to ESP. Certain secondary hypotheses concerning the nature of ESP were also investigated.

This study is based in part upon a thesis submitted by the author to the Department of Social Relations of Harvard University, in partial fulfillment of the requirements for the degree of Bachelor of Arts.

Studies in Extrasensory Perception

Experiments Utilizing an Electronic Scoring Device

S DAVID KAHN

Introduction

Since we are concerned here with experiments in extrasensory perception, we shall begin by defining the term as we use it: "When a large number of observations is made, and a statistically significant correlation exists between a specified set of completely random events and symbolic behavior *intended to refer to those events*, ESP by our definition is present. These observations must be made under the following condition: no information shall be permitted to pass between the target material or other source of relevant information and the subject by means of any energy process to which the human organism is known to be able to react."

In the present paper only the "clairvoyant" form of ESP is studied, in which a physical object, and not the content of another person's mind, is used as the target.

The primary problem which prompted the following research dealt with this question: "Can the phenomena defined as extrasensory perception be observed under conditions which eliminate all counter-hypotheses, regardless of how improbable these counter-hypotheses may seem?" In other words, can we approach a limit of experimental rigor which will satisfy the most critically-minded observers? It should be emphasized that the "scientific method" does not *logically guarantee* the validity of its results; ultimately, all conclusions are based upon certain criteria of significance based in turn upon personal judgment. Fortunately, there are certain standards which are generally agreed upon by the large majority of investigators in various fields of research. An example of this is the mathematics of probability to be used when the question arises whether a phenomenon can be attributed to chance. It should be realized, however, that this agreement is affected by the importance of the phenomena being described. Moreover, when a proposition causes little interest, or seems to be very reasonable on the basis of previous observations, it usually has few critics, the principle apparently being that if the hypothesis is wrong, this fact will be discovered before much damage has been done. It is in the case of observations which have far-reaching significance that the prerogative of individual judgment is exercised most vigorously. For example, one might here consider the Lamarckian controversy, or the controversy over the wave and particle theories

of light which raged between Newton and Huygens. This is also the case with parapsychology. Even the most radical critics admit the large-scale theoretical problems which would be evoked by the ESP hypothesis should it prove to be valid. Thus, to the present writer, it would seem that the burden of proof is on the parapsychologist, in the sense that intelligent counter-hypotheses, even if considered in his *personal* judgment to be highly improbable, must be dealt with. For in the present state of scientific method it is all too easy to observe that agreement simply will not be reached by *arguing the probability* of a counter-hypothesis. It must be considered experimentally, by actually eliminating all reasonable possibility of the counter-hypothesis. This, then, is what this research has primarily attempted to accomplish.

It was also felt that it would be extremely valuable to test simultaneously a number of the secondary hypotheses concerning the nature of ESP which have been offered in the literature. For unless a hypothesis is such that it can be proved false, if it *is* false, it can be of no value to scientific progress. This can only be done when a hypothesis is offered which attempts to predict the occurrence of future phenomena with an accuracy approaching the probability limit of one: absolute certainty. It is only at this point that one can manipulate variables until one of them is found systematically to alter the phenomena, thus bringing under control a variable which can be considered crucial. It would seem that parapsychology may be close to having achieved a repeatable experiment, in the verification of some of the secondary hypotheses that have been offered. However, we can only be sure of this by testing them sufficiently, and by *independent* investigations. Therefore, it is almost in a sense the "scientific obligation" of investigators doing research in a given field to attempt to verify its hypotheses until they are either accepted as valid or have been demonstrated to be invalid. This, then, was the secondary purpose of this research.

In the experimental work to be reported six hypotheses were tested. They can be classified into two broad areas on the basis of whether they are concerned with differences among the individuals being tested, or whether they are concerned with the group as a unit, being tested under varying conditions. Below, for the sake of clarity, we have presented the hypotheses, with citations of some of the relevant literature in parentheses, but without comment. We shall then take up each one separately, discussing in turn the methods used in testing the hypothesis, the results, and the conclusions. They are as follows:

I. Primary Hypothesis.

1. ESP can be demonstrated by a methodology which excludes all possible counter-hypotheses (11, 16, 21, 22).

The method must offer the possibility of being independently and precisely repeated.

II. Secondary Hypotheses.

- A. Hypotheses whose conditions ignore individual differences.
 2. A gross decline effect will be observed (6, 30, 38).
 3. ESP can be demonstrated under conditions in which considerable distances separate the target from the subject (15, 17).
 4. The freer the psychological conditions under which the subject operates in ESP tests, the greater will be the number of direct hits on the target (1, 8, 9, 10, 18, 23).
- B. Hypotheses whose conditions specify individual differences.
 5. The attitudes which a subject brings with him into the ESP situation will be found to correlate with his ESP ability (1, 3, 8, 9, 10, 19, 23-33, 39).
 - a. "Sheep" (those who accept the theoretical possibility of ESP) will tend to score better than chance expectation, and "goats" (those who reject this theoretical possibility) will tend to score at or below chance expectation (24-33, 39).
 - b. The greater the degree of acceptance of the actual occurrence of ESP, the higher the scoring level will be (1, 31).
 - c. Individuals who would be pleased at demonstrating ESP will tend to score higher than those who would be indifferent or displeased.
 - d. Subjects who show confidence in their ESP ability will tend to score higher than those subjects who do not show such confidence (1, 18, 24-27).
 - e. "Expansive" subjects (as defined below) will tend to score above chance on direct hits in clairvoyance type ESP tests while "compressive" subjects (as defined below) will tend to score at or below chance (8, 9, 10).
 - f. Subjects who report they would be disturbed by the possession of ESP ability will tend to score below chance expectation, while those who report they would not be disturbed will tend to score above chance expectation.
 6. Subjects who are rated as well adjusted will tend to score above chance expectation, while subjects who are rated as poorly adjusted will tend to score below chance expectation (7, 28, 29, 31).

Hypothesis I

ESP can be demonstrated by a method which excludes all possible counter-hypotheses. The method must offer the possibility of being independently and precisely repeated.

Discussion and Methods

We might list as follows the necessary conditions for a theoretically adequate method of eliminating all counter-hypotheses to ESP:

- a. There must be no theoretical possibility of sensory knowledge or inference of the target by the subjects,
- b. There must be no theoretical possibility of error in the comparison of the target and the subjects' attempts to duplicate the target,
- c. If an instrument is used to make such comparisons, there must be the possibility of independently and mechanically checking the accuracy of the instrument,
- d. The statistical analysis must be appropriate to the problem at hand,
- e. There must be no loss or selection of data.

On the basis of these requirements it was felt that the crucial problem consisted of finding a machine for scoring the tests which could be demonstrated to be reliable. After investigation, it was decided that the International Test Scoring Machine, manufactured by the International Business Machines Corporation (IBM), could be used for this purpose.

A testing technique was then developed and utilized in testing five groups of subjects. In Series 1, 3, 4 and 5, the standard "answer sheet" IBM form I.T.S. 1100 B 107 (see reproduction in Appendix II) was used both as target and record sheet. On each side of this answer sheet are 150 calls or sets, each offering five choices. The sets are numbered 1 - 300, and are arranged in columns of 30. Each set consists of five small vertical rectangles arranged in a horizontal row, thus:

1	2	3	4	5
..
..
..
..
..

One of these sheets was prepared as target. Using tables of random numbers (14) to make the selections, *one* rectangle in *each* of the 300 sets was blacked in. Thus the target sheet consisted of 300 individual targets, randomly arranged. Randomization prevented the theoretical possibility of rational inference regarding the target.

In Series 2, IBM form I.T.S. 1000 B 106 was used. This differs from the other form in that it has space for only 150 calls, the reverse side of the sheet being blank.

Except in Series 2, no effort was made to insure that an equal number of targets would appear in each position (rectangles 1 through 5). In Series 2 an equal number of rectangles was filled in for each of the five possible target positions. The equalization was accomplished by disregarding the number in the random number tables representing a particular position as soon as that position was represented fifteen times, within each half of the target series (that is, from targets 1 to 75, and 76 to 150). Thus, each target appeared fifteen times in a given half of the target record.¹

The tests were administered to five groups of subjects, a group consisting of from nineteen to sixty-three subjects. Each subject was given an answer sheet and was informed of the nature of the target sheet. He was either told or shown its general location.

In Series 1 (May, 1949), the target was locked up in a file cabinet placed in the Department of Psychology at Harvard. The subjects did not know the exact location. In Series 2 (June, 1949), the target was in a sealed envelope in a locked room. In Series 3 and 4 (March, 1950), it was in the room with subjects, between pieces of cardboard and in a heavy sealed envelope. In Series 5 (March, 1950), it was in a locked metal box; the subjects had seen the box and knew its general location.

In all series the subject was asked to duplicate the target sheet

¹ It was pointed out by statisticians at Duke University that the method used for preparing this particular target sheet did not produce a random order of events due to the restrictions imposed toward the end of each half of the target sheet. As the analyses usually applied in such situations are based on the assumption of "randomness" throughout the target series the method raised questions that could be answered only by further study of the data. If the subject showed preference for certain response positions, these might operate to give spuriously high or low scores in the nonrandom part of the target sheet.

In a recheck at Duke it was found that positions 2, 3 and 4 were called much more often than the first and fifth. In view of the unusual way in which the target order was prepared, it seemed safer to evaluate the truly random selections of the target sheet separately. The random sections of the data (113 of the 150 positions) were re-evaluated by the Greville method and a CR of 2.24 was obtained. When the entire data were evaluated by the Greville method, a CR of 1.74 was obtained. It is therefore evident that in this particular series, the nonrandom character of part of the target did not produce spuriously high results, but, if anything, spuriously low results.

as nearly as possible. No specific instructions were given as to how this should be accomplished. It was merely stated that some subjects liked to try to visualize targets, while others preferred to fill out the sheet by following their "hunches." They were told to follow whatever system seemed best to them, although never to forget what they were trying to accomplish.

The subjects were all volunteers with the exception of the last group, Series 5; these were paid subjects being used in an investigation of projective tests performed by a department of Harvard University. They were requested to take the tests, and were remunerated for their services.

Series 5 actually served two purposes. In itself it was intended as an experiment comparable to the other series reported in this paper, and as a check on several of the hypotheses discussed here, which had already been formulated at that time. In addition, it was the first of a set of six IBM tests performed with the same group of subjects in order to study such decline effects as might occur over a long series, and to investigate correlations of ESP with certain personality variables. A preliminary report of this work forms Appendix I of these *Proceedings*. A full report does not belong here because results from a protracted series constitute a problem quite different from the problem of this report. The basic hypothesis of this type of decline-effect work is necessarily contradictory to the fundamental idea of the work reported here, where the objective was to elicit positive ESP scoring.

All subjects were college students, most of them male, with the exception of Series 2; here the subjects were sixty-three adults, each sex being represented about equally.

There was no selection of subjects in terms of ESP ability, except in Series 1. Here, only subjects who had shown a deviation of at least +3 on 250 calls in a previous experiment using ESP cards were used. Although this deviation is certainly quite insignificant, it was nevertheless felt that there would be more positively scoring subjects in such a selected group than would be found in a random sample.

At the end of each series, the sheets were collected by the experimenters, and deposited with either the Harvard Bureau of Tests, or the Educational Records Bureau in New York City. They were scored by the IBM machine mentioned above. It operates by recording the closures of electrical circuits, each closure facilitated by the pencil mark made by the subject, if it has been placed in the correct position. If by chance, the pencil mark is not heavy enough, and the hand inspectors do not correct this, the mark will not be counted.

The sheets were hand inspected before being scored. This inspection had another function besides that of darkening faint marks. Errors of omission (sets where the subject has made no mark) and errors of commission (where the subject has made two or more marks) were counted, since the machine does not detect them. Such sets are not included in the totals. Undiscovered errors of omission would lead to spuriously low scores by increasing the number of apparent calls made; undiscovered errors of commission would lead to spuriously high scores by raising the probability of a hit from one-fifth to two-fifths or more. Such corrections account for the odd numbers of calls and fractional expected hits that appear in the tables.

[In Series 1, 2 and 5, the distance between target and subjects ranged from under one mile to over 500 miles. In Series 3 and 4 the target sheets were present in the same room with the subjects. However, they were always placed between two heavy sheets of cardboard and then sealed in heavy manila envelopes, under the constant surveillance of two experimenters.]

No targets were removed from their safes or envelopes until all tests were collected. Without comparisons, they were deposited in the testing offices. No loss or selection of data could take place since the testing offices made records of the number of sheets they corrected. The accuracy of the machines is tested by periodic hand checks of data and by other similar machines to give cross checks.

All data were scored twice. With this procedure, we may assume that random errors by the highly reliable IBM machine or its operators are very small. These errors may be uncovered by a detailed handcheck (cf. footnote, p. 9). The crucial advantage of the IBM technique is the *independent* check of the results which it provides.

The last important problem to be discussed concerns the statistical methods used to evaluate the data. The usual method of evaluating ESP data is based upon the binomial formula (20). Since, however, the results of these series are based on tests in which many subjects made calls for the same targets, it must be recognized that the binomial method of evaluating data may not be appropriate. If there were any group tendency among the subjects that led them to mark the *same choices on the same trials*, the binomial formula would overestimate the significance of the results. A method for evaluating multiple calling data of this sort without making any assumption regarding independence among the responses has been presented by Greville (4).

The Parapsychology Laboratory at Duke University kindly offered to analyze the data of the first two series by applying the Greville method under the direction of Dr. T. N. E. Greville, and by doing

a hand tabulation of the data under the supervision of Dr. J. G. Pratt. After correction for blank and double guesses, the hand tabulation agreed with the tabulation given by the machine (11). As will be observed later, the Greville and binomial formulas yielded approximately the same CR, thus demonstrating that there were no significant group preferences operating in these two series that would yield spurious deviation by analysis based on the binomial formula.

Since these first two series when combined were significant, it was felt that there was no need to apply this extremely cumbersome technique on further series. Moreover, it was felt that in five series of tests, using six target sheets which included eleven sides, any preferential pattern would be averaged out. However, the writer would rest his case on the empirical observation, based on the application of both the Greville and binomial technique to the first two series, that group preferences were not yielding spuriously high results.

Results

The results of the five series are presented below in Table 1. The deviation refers to the number of hits made by subjects above the number expected by probability theory on the basis of chance alone. The percentage column represents the deviation divided by the hits expected by chance, in per cent terms. It is presented in this table, and the ones to follow only in order that the observer may have a ready means of comparing the various cuts made through the total data in terms of percentage increase in scoring over that expected. The CR column is used to compare these deviations when size of sample is to be taken into consideration. The critical ratios are of course comparable in terms of the probability which they represent.

Figure 1 presents the percentage deviation and size of sample in terms of a graph, to facilitate comparisons between the five series, when these comparisons are specifically warranted in the discussions.

As is evident in Table 1, a deviation as large as the one observed should be expected to appear only once in about 2000 experiments¹ of similar length.

It should be noted also that the first two series² are independently significant, when taken together, either by the binomial formula (p of .003) or by the Greville method (p of .006).

¹ Since our predictions in this experiment have stated *direction* of deviation, only the area under half the binomial curve has been used to estimate probabilities. In those cases where no prediction as to the direction of deviation has been made, the area under the full curve is used.

² A more extensive report of these two series can be found in the *Journal of Parapsychology*, 1949, 13, 177-185.

TABLE 1¹
TOTAL RESULTS (ALL SERIES)

Series	Number of Subjects Registered	Total Calls Registered	Deviation	Excess of Deviation over Expectation, in %	Binomial		Greville	
					CR	p	CR	p
1.	21	6278	+ 63.4	5.05%	2.00	.023	1.85	.032
2.	63	9399	+ 75.2	4.00%	1.94	.026	1.74	.041
3.	22	6554	+ 44.2	3.37%	1.36	.087		
4.	52	15360	+ 10.0	0.33%	<1	—		
5.	19	5687	+ 78.6	6.91%	2.61	.005		
Totals	177	43278	+271.4	3.14%	3.26	.00056	✓	

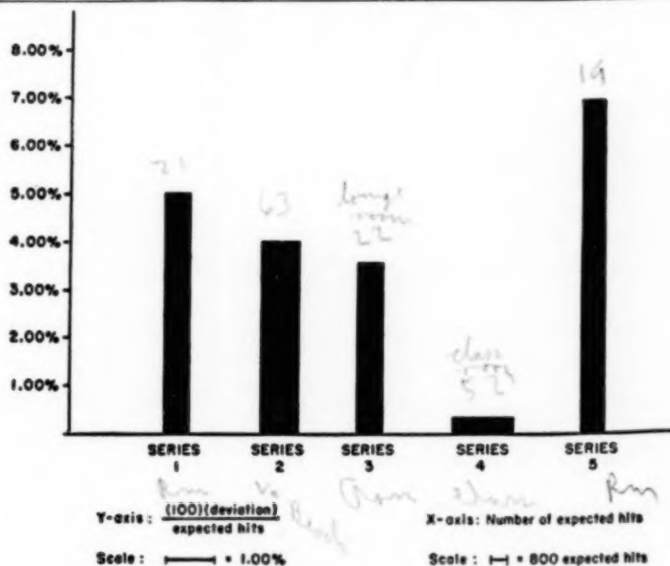


FIGURE 1

Deviation from expectation, in terms of percentage, for all five series.

¹ Subsequent to the submission of the manuscript, the data were thoroughly handchecked by Mrs. L. A. Dale, Research Associate of the Society. An error of four hits in the total count by the machine was found, and six previously unnoted errors of omission or commission were discovered. These errors constitute .04% of the number of hits and .2% of the number of calls. They have no effect on the significances reported.

Since the object of this research was to eliminate as far as possible any theoretical possibility of counter-hypothesis, the question of errors on the part of the subject, which had to be recorded by hand, is an important one. The breakdown of hand-recorded errors is as follows:

Errors of omission, where subject failed to make any mark	96
Errors of commission, where subject made two marks	50
Errors of commission, where subject made three marks	1

(Omitted calls due to failure to finish test by four subjects in Series 3 and 4 . . . 225).

As was pointed out in an earlier section, these errors were noted by scanning each sheet by hand. This is the standard procedure in the use of the IBM Test Scoring Machine.

We shall now consider how errors on the part of the hand inspector may have yielded spurious results. Let us assume that the inspector misses ten per cent of the actual marking errors, and that the errors he has overlooked are distributed like those he has found. Then, in the present data, he would miss ten errors of omission and five of commission. Calculation shows that this would lower the expected hit total by two and the actual hit total by an amount whose most probable value is one. Needless to say, this would not affect the significance of the results. Actually, of course, it is unlikely that a checking mistake of this magnitude would occur. (Cf. footnote, page 9, for data on this point.)

The possibility of errors by the machine itself or its operators has already been mentioned, and it has been pointed out that these can be expected to be minute. Their actual magnitude can be estimated by an application of the most theoretically perfect checking method: by following the machine check (and its accompanying hand count of omissions and commissions) by an entirely independent hand check of all of the data. Comparison of the results of these two checking techniques would permit the accurate localization and correction of any discrepancies. This "idealized" checking method was applied to Series 1 and 2 by the Duke Laboratory of Parapsychology, as already reported, and no discrepancies were found. Hence a first estimate of the machine error would be zero. Furthermore, this last check was later applied to the remainder of data; the discrepancies observed are reported in the footnote to Table 1, page 9.

Conclusions

In view of the highly significant deviation from chance expectation it can be concluded that the results in Table 1 are non-chance phe-

nomena. Furthermore, it is the writer's opinion that no counter-hypotheses can be offered to invalidate the conclusion that extrasensory perception, as defined in an earlier section, was operating in the group of 177 subjects.

It should be observed that this experiment using IBM targets differs qualitatively from similar experiments with cards. The subjects were asked to intuit the spatial location of marks on a page, in contrast to the task in a card experiment. In the latter case, the subjects are asked to make successive differentiations between different kinds of geometrical symbols. Here only one symbol, a black mark almost entirely devoid of any affective connotation, is used. At this point, it is impossible, of course, to tell whether these target distinctions are related to the paranormal functioning of the subjects.

The results are particularly interesting in the light of the experimental work carried on at Duke. First, it has again been possible to repeat the findings of Rhine and his associates entirely independently. Secondly, tightening of experimental controls does not preclude the achieving of results that are quite as significant as those of experiments carried on under looser experimental controls.

In the writer's opinion, the evidence for extrasensory perception reported here is of particular value because of the mechanical scoring technique employed. In the first place, the percentage of error in the machine totals, as revealed by a subsequent hand check, is extremely small; hence the machine alone may be used with confidence in future experiments. More important, however, the IBM method provides a unique opportunity to attack the scoring error problem directly. The method makes it possible to apply two completely independent checking techniques (human and mechanical scoring) to the same data, and thus permits an estimate of the error in either one. The close agreement between the two methods on the present data indicates that both are highly reliable. As already suggested, a still higher level of accuracy, if required, can be achieved by using both and rechecking any discrepancies.

No scoring system can ever be completely free from error. The theoretical perfection described earlier as the goal of this or any other method can never be completely attained. Nevertheless, it is felt that the techniques used here have reduced the inherent error to a level that should not disturb even the most rigorous critic.

It should be fully noted that considering the large number of IBM machines which are easily available, the method offered above is quite practical in the sense that any investigator who cares to can conveniently attempt independent repetitions, the importance of which we have noted earlier. And, lastly, all scoring will thereby be done by disinterested parties.

Hypothesis II

A gross decline effect will be observed.

Methods and Results

One of the commonest effects reported in the literature is a decline in scoring as the test progresses. From the periodic U-curves found in the card-type experiments, it was suggested that the effects, in these cases at least, were probably psychological in nature. It appears from the literature that the U-curves are most commonly found in the type of card experiment in which the cards are called by the subject down through the pack, the pack being stacked and untouched during the experiment. Since the IBM sheet was not analogous to this particular type of situation, nothing could be predicted about the appearance of U-curves in this experiment. It was therefore felt that the only thing we could definitely expect to find would be a decline effect, and that this could be observed most easily by comparing one side of the answer sheet with the other, expecting that there would be more hits on the first side than on the second. (It was found that there was no noticeable decline effect in Series 1 within one side of the sheet; i.e., when calls 1-75 were compared with calls 76-150, or when calls 151-225 were compared with calls 226-300.)

As will be seen from Table 2, there was a decline *between one side of the sheet and the other* in Series I. It was then decided that in Series 2, only one side of the sheet would be used. This was done with the following consideration in mind: if the gross decline effect in Series 1 was purely psychological, we might expect it to have occurred in terms of the subjects' structuring of the task. Therefore, if we shortened the test to merely one side of the sheet, then perhaps within this single side a similar decline effect would be observed. If it were not observed, then we might suspect that the effect was not entirely psychological.

When this analysis was made of Series 2, it was found that there was a slight, although insignificant, *incline* effect present when calls 1-75 were compared with calls 76-150.

The next three Series: 3, 4, and 5, were carried out with two-sided sheets. This was done to see whether the *revised* hypothesis, that there would be no decline effect within the page, but that there would be a gross decline effect between side 1 and side 2, would be borne out. A detailed analysis of all position effects has not yet been made with a special IBM machine designed for this particular purpose. As can be seen from Table 2, the gross decline effect comparing

side 1 and side 2 of each sheet consistently appeared in these three verification series. The comparison of side 1 of all series with side 2 of all series is presented below in Table 2.

TABLE 2
DECLINE EFFECT FROM SIDE 1 TO SIDE 2

Series	Number of Subjects	Total Calls Registered	Deviation	Excess of Deviation over Expectation, in %	CR
<i>Side 1</i> <i>Calls 1-150</i>					
1.	21	3142	+ 64.6	10.28%	2.88
2.	63	9399	+ 75.2	4.00%	1.94
3.	22	3284	+ 35.2	5.36%	1.54
4.	52	7757	+ 47.6	3.07%	1.35
5.	19	2845	+ 51.0	8.96%	2.39
Totals	177	26427	+273.6	5.18%	4.21
<i>Side 2</i> <i>Calls 151-300</i>					
1.	21	3136	- 1.2	-0.19%	<1
2.	—	—	—	—	—
3.	22	3270	+ 9.0	1.38%	<1
4.	52	7603	- 37.6	-2.47%	<1
5.	19	2842	+ 27.6	4.86%	1.29
Totals	114	16851	- 2.2	-0.07%	<1

Side 1
CR = 4.21
p = .000,013

Side 2
CR = <1
p = —

Difference
CR_{diff} = 2.66
p_{diff} = .004 ✓

As can be observed, the total positive deviation on side 1 is so high that it would be expected to occur by chance only about once in fifty thousand experiments of similar length. The obtained difference between the means of side 1 and side 2 should occur by chance approximately once in two hundred and fifty similar experiments. This type of consistent effect can be considered to support the basic ESP hypothesis.

Conclusions

✓ We can say at this point only that the significant scoring took place entirely on the first side of the combined sheets. As can be seen from Table 2, the average score for the second side of the combined sheets was almost chance expectancy.

We cannot fail to note the highly significant critical ratio, 4.21, which was obtained by combining the first sides of all the sheets. The difference between the means of the two sides is also significant at the .004 level of confidence.

This is a very interesting effect, and quite consistent with the general hypothesis. However, more detailed position analyses will be necessary to determine whether there were any incline or decline effects *within a single side* of the sheet when considered alone.

Hypothesis III

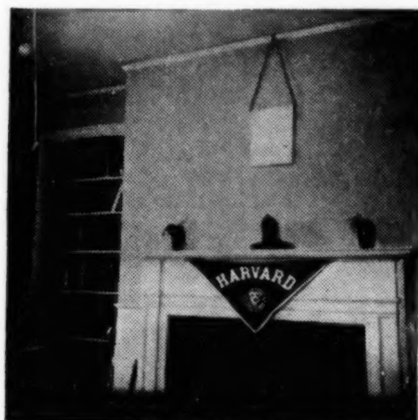
ESP can be demonstrated under conditions in which considerable distances separate the target from the subject.

Methods and Results

✓ Three series of subjects were tested in a situation in which the target was placed at varying distances from the point at which they were taking the test. In Series 1, the target was prepared by a member of the staff of the Psychological Laboratories, and then locked in a filing cabinet in his office in Memorial Hall at Harvard. Twenty-one subjects were given IBM answer sheets. They were then informed of the nature of the target sheet, and that it had been placed in a filing cabinet in the Laboratories in the basement of Memorial Hall. They were told to fill out the sheets in their rooms at their leisure, and "when they felt like it."

In Series 5, the same procedure was followed with nineteen subjects. In this case, however, the target was placed in a small locked file box. Each subject was shown the box and its location before he was given the answer sheet to fill out. The distances concerned here naturally vary from subject to subject; however, they can be considered to fall within the range of from one-half mile to about two miles.

In Series 2, the same general procedure was followed. In this case, the target was hung in a sealed envelope on the wall of a locked room in Lowell House. Sixty-three subjects were shown a photograph of the room and building, with the envelope (empty) in position (see below). The subjects were all located in Virginia Beach, Virginia.



The distances between the target and subjects was therefore approximately five hundred air miles. It should be borne in mind that these three groups cannot be compared with one another since they were drawn from different populations, as pointed out earlier. No experimenter who was aware of the order of the targets had any contact with any of the subjects. Thus there was no theoretical possibility of unconscious sensory cues being transmitted to the subjects.

TABLE 3
RESULTS OF EXPERIMENTS INVOLVING
SUBSTANTIAL DISTANCES (SERIES 1, 2, 5)

Series	Number of Subjects	Total Calls Registered	Deviation	Excess of Deviation over Expectation, in %	Binomial		Greville	
					CR	p	CR	p
1.	21	6278	+ 63.4	5.05%	2.00	.023	1.85	.032
2.	63	9399	+ 75.2	4.00%	1.94	.026	1.74	.041
5.	19	5687	+ 78.6	6.91%	2.61	.005		
Totals	103	21364	+217.2	5.08%	3.71	.0001		

As can be observed from Table 3, the deviations in Series 1 and 2 approach significance, while the deviation in Series 5 is independently significant. As pointed out earlier, Series 1 and 2 combined are significant either by the binomial (p of .003) or Greville (p of .006) formula. The total deviation of the three series combined should be expected to occur once in ten thousand experiments of similar length by the chance hypothesis.

Conclusions

On the basis of the highly significant deviation observed in these three series, it would seem quite safe to conclude that ESP was operating. Secondly, we can conclude that ESP can operate over considerable distances. In the writer's opinion, this does not warrant any speculation concerning the so-called "non-physical" nature of ESP. All we can safely say is that the range over which ESP can operate is greater than the range over which sensory perception has been observed to operate, given the nature of the target and the intensity of the physical energy transfers involved.

Probably the most important point to be made in considering the results of these three series is their bearing on the primary ESP hypothesis. By no stretch of the imagination could sensory cues be considered present in these tests. As pointed out earlier, this type of counter-hypothesis to ESP has been a common one in the critical literature of parapsychology. It is these three series which the writer would consider as best meeting any possible counter-hypotheses to ESP, however problematic.

Hypothesis IV

The freer the psychological conditions under which the subjects operate in ESP tests, the greater will be the number of direct hits on the targets.

Methods and Results

A great deal has been said as to the situational conditions which are most favorable for the operation of ESP. However, the vast majority of these observations are clinical rather than experimental in nature. It might be possible to summarize a good proportion of these clinical observations in terms of the relative rigidity or freedom under which the subject operates in the experimental situation. Unfortunately, there have been only a very few attempts to define these terms operationally. It will be recalled that Scherer (23) showed that calls made at will, spontaneously, gave a significant number of hits, while calls made on request of the experimenter resulted in scores at the level of chance expectation. In Schmeidler's early work, the "goats" were tested under unpleasant conditions, while the "sheep" were not, although it is impossible to tell whether the goats' low scores were due to their attitude towards ESP or to their working conditions or both.

It was therefore decided that it would be worth while to attempt to define operationally two situations which approached most nearly the clinical observations as to "good" and "bad" conditions for ESP. There was no attempt to decrease scores artificially by being unfriendly to the subjects; it was merely a matter of setting up two situations and then trying to get as much ESP out of the subjects as possible under each condition.

The first type of situation was that present in the three distance experiments. Here, subjects were asked to take part in ESP tests. It will be recalled that in Series 1 the subjects were selected on the basis of previous tests in which they had shown positive though insignificant deviations. Series 5 consisted of the subjects who were requested to take part in the experiment, and who were remunerated. Subjects in Series 2 were volunteers and were all favorably inclined as to the possibility of ESP.

In the "free" situation, the procedure was as follows: an experimenter explained the nature of the test to the subject. He then gave him an answer sheet, instructing him to attempt to duplicate the target sheet during the next few days. He was told to pick a time when he felt relaxed, undisturbed, and when he could be free from outside

disturbances. He was to fill it out whenever he felt like it. He was told that he did not have to fill it out all at one sitting; if he became bored, he could finish it at some later time.

The "rigid" situation was operationally defined as group testing. In Series 3 a group of Massachusetts Institute of Technology students were used. They were members of a group interested in parapsychology. The test was administered to the group in a lounge room, with the subjects sitting informally about in armchairs or about a seminar-type table. They were orally told of the nature of the experiment. They then were given an attitude questionnaire to fill out, and then asked to do the experiment. Most of them knew one another, and the session seemed quite informal and easy.

Series 4 was also a group experiment. Here subjects were asked to volunteer by means of advertisement. The testing was done on two days, a different target sheet being used each day. The tests were conducted in a large classroom. The subjects were given printed directions, which were read aloud by the experimenter. In addition to the attitude questionnaire, and the ESP tests themselves, a lengthy personality inventory was used. Despite the attempts of the experimenters, the sessions turned out to be quite long, and quite tiring. The sessions seemed to tend toward extreme formality, and there was no joking or bantering as in Series 3.

On the average, the subjects used in the "free" situation are fairly comparable to the subjects used in the "rigid" situation. However, it must be borne in mind that what differences there were might actually be critical ones in terms of ESP, and that, therefore, the results must be considered in the light of this fact.

TABLE 4
COMPARISON OF FREE AND RIGID SITUATIONS

	Number of Subjects	Total Calls Registered	Deviation	Excess of Deviation over Expectation, in %	CR	p
Free Situation (Series 1, 2, 5)	103	21364	+217.2	5.08%	3.71	.0001
Rigid Situation (Series 3, 4)	74	21914	+ 54.2	1.24%	<1	—

Difference

$CR_{diff} = 2.00$

$p_{diff} = .023$

If we take the free situation and compare it to the *extreme* rigid situation (Series 4), the differences are even greater.

TABLE 5
COMPARISON OF FREE AND EXTREME
RIGID SITUATIONS

	Number of Subjects	Total Calls Registered	Deviation	Excess of Deviation over Expectation, in %	CR	p
Free Situation (Series 1, 2, 5)	103	21364	+217.2	5.08%	3.71	.0001
Extreme Rigid Situation (Series 4)	52	15360	+ 10.0	0.33%	<1	—

Difference

$$CR_{diff} = 2.24$$

$$p_{diff} = .012$$

As can be seen from Table 4, the scoring levels of the "free" and the "rigid" groups are significantly different from one another at about the 2 per cent level of significance. When the *extreme* groups are compared in Table 5, the difference is significant at close to the 1 per cent level.

Conclusions

The results of Table 4 would seem to support the original hypothesis. Since the subjects were different in the two situations, however, we can only consider this conclusion as tentative. The attitudes among the two groups were on the whole fairly similar, in terms of how favorably inclined they were to the ESP hypothesis, but the subjects in Series 1 were selected as described above, and the subjects in Series 5 were remunerated.

It should be noted that while the breakdown in Table 5, in which the more informal group test is omitted (Series 3), may be interesting, it is nevertheless not fully justified, since there is no clear operational method of differentiating between this less rigid group situation and the more rigid group situation of Series 4. It was done merely by clinical observation, which is unsatisfactory for our purposes.

In conclusion, we might say that on the basis of these results it would seem very well worth while to evaluate the original hypothesis

clearly by carrying out a similar experiment in which the same subjects were used in both the group and the free situations.

Since it will unfortunately be a long time before we can hope adequately to measure and then isolate the variables introduced into ESP tests *by the experimenters*, it might be profitable to eliminate these variables as far as possible. In the ESP situation, this can be done by eliminating as far as possible their source, namely, contact between experimenter and subjects. One might therefore consider the free situation as defined above in the light of the effect of the experimenter on the subject. Here the experimenter-subject relationship consists in doing little more than explaining the test and giving the subject the answer sheet. It might well be that such a short sequence gives these field effects little chance to operate. If this were the case, then we could consider experimenter-introduced variables present to a minimum degree. It would then probably be a much easier job to measure the more clear-cut attitude and motivational variables that remain, with an eye to accurate prediction of scores under conditions which would be standardized for all experiments and subjects to a higher degree.

Hypothesis V

The attitudes which a subject brings with him into the ESP situation will be found to correlate with his ESP ability.

1. *Discussion and General Methods*

There have been several specific studies correlating attitudes and ESP scores (1, 3, 8, 9, 10, 12, 24-27, 29, 33). In addition, there are the numerous clinical observations on the basis of which experimenters have tried to specify the most favorable attitudes for the occurrence of demonstrable ESP. Therefore it was decided first to attempt to duplicate several of the important previous experimental studies of attitude, to the extent that they were clearly enough described in the literature to permit repetition. These included Schmeidler's hypothesis that those who accept the theoretical possibility of ESP (the "sheep") will score higher in clairvoyance than those who reject it (the "goats") and Humphrey's hypothesis that those who make "expansive" drawings will score higher in clairvoyance than those who make "compressive" drawings.

An attitude questionnaire was made up (see appendix III) and given to the seventy-four subjects in the two group series, 3 and 4. All subjects were first asked to fill out these sheets. They were then requested, before taking the IBM test, to attempt to duplicate a picture which was contained in a manila envelope in the front of the testing room. All that was desired from this was an index of the compressiveness or expansiveness of the subjects' drawings; however, it was felt that they should believe their drawings were a part of the test of ESP in order that the drawings would better express their attitude in the specific ESP situation, as differentiated from their attitude to the general experimental situation. For purposes of clarity, we shall consider each of the six attitudes which were indexed separately.

2. *Subhypotheses*

- a. "Sheep" will tend to score better than chance expectation, and "goats" will tend to score at or below chance expectation (1, 24, 25, 26, 27, 33).

Only those who *entirely reject the possibility* of ESP are goats by our criterion. To isolate these individuals, the following questions were asked:

Do you think that extrasensory perception is theoretically possible:

I. In this particular experiment?	Yes	62	(number of responses)
	No	12	
II. Under other circumstances?	Yes	70	
	No	4	

The result of the breakdown on the basis of this question is presented in Table 6. No subjects answered "No" to II and "Yes" to I. Only four subjects out of seventy-four answered "No" to both questions.

As can be seen from Table 6 those subjects who entirely rejected the idea of ESP, and there were only four of them, averaged below chance in the predicted direction. The sheep scored above chance, although not significantly.

TABLE 6
EFFECT OF BELIEF IN THE POSSIBILITY
OF ESP ON ESP SCORE (SERIES 3, 4)

Those Believing ESP is —	Number of Subjects	Total Calls Registered	Deviation	Excess of Deviation over Expectation, in %	CR	p
"Impossible here only"	8	2387	+21.6	4.52%	1.1	—
"Impossible anywhere"	4	1197	— 9.4	—3.92%	<1	—
"Possible here and elsewhere"	62	18330	+42.0	1.15%	<1	—

There are probably two reasons why we got so few goats in comparison to Schmeidler who usually gets about half and half. In the populations we used there were probably few individuals who were willing to be dogmatically certain about the impossibility of anything, regardless of how improbable they may have considered it to be. Schmeidler's populations may have differed in this general respect. Secondly, Schmeidler's populations are usually drawn from classes, where goats are required to participate. The fact that in our case the subjects were volunteers may have acted to select out the goats, who presumably do not usually take the time to participate in ESP experiments of their own free will.

In this experiment a completely standard criterion and method was used to divide the sheep from the goats. In most of Schmeidler's studies the subjects rated themselves, as here, but only after Schmeidler finished explaining to them what the differences are between a

sheep and a goat. One can only deduce from the literature what she actually says to these subjects, her descriptions in the experimental reports being of a general nature only. The present writer feels, from his impressions of Schmeidler's work, that she would have found more goats in this particular population than he did, although he used what he felt to be the criterion closest to the one she actually uses.

Our conclusion from this is quite simple. As Schmeidler's work now stands in the published literature, it is extremely valuable, but, using her reports as a methodological guide, difficult to repeat rigorously. Further clarification of her methods, to the point where an independent investigator can attempt to repeat them precisely, would be valuable.

- b. The greater the degree of acceptance of the actual occurrence of ESP, the higher the scoring level will be.

It was not expected that belief in the theoretical possibility of ESP would differentiate the groups adequately. Furthermore, it was felt that what was really concerned here was not intellectual judgment *per se* but the emotional set which the individual brought with him to the experiment. Both Schmeidler (31) and Bevan (1) had attempted to differentiate between their sheep on the basis of strength of belief. Schmeidler found the "sheep-plus" (definite believers) scored higher than the "sheep-minus" (uncertain believers). Bevan found that the undecided sheep scored better than the sheep who were positive believers; but neither result is statistically significant. We hoped to throw some light on this question by asking the following:

Do you think it probable that extrasensory perception actually ever occurs?

(a) Very probable	31
(b) Fairly probable	23
(c) Slightly probable	15
(d) Improbable	4
(e) Very improbable	1

By considering *c*, *d*, and *e* together, we can split the group into three ranges of belief. The scoring levels are presented in Table 7.

It can be seen from Table 7 that none of the scoring trends is significant. Numerically, although not statistically, they can be said to support Bevan's findings rather than Schmeidler's. In other words, the group that believed most strongly in ESP did not score higher than the less strong believers; and in this case actually scored lower. It would seem likely that the differences reported by Schmeidler,

Bevan, and the present author, while theoretically due to the differential effect of different strengths of belief in the occurrence of ESP, are in actual fact mere chance fluctuations.

TABLE 7
EFFECT OF BELIEF IN THE PROBABILITY
OF ESP ON ESP SCORE (SERIES 3, 4)

Those Believing ESP is —	Number of Subjects	Total Calls Registered	Deviation	Excess of Deviation over Expectation, in %	CR	p
"Very Probable" (a)	31	9088	+17.4	0.96%	<1	—
"Fairly Probable" (b)	23	6841	+44.8	3.27%	1.35	.09
Others (c, d, e)	20	5985	— 8.0	—0.67%	<1	—

- c. Individuals who would be pleased at demonstrating ESP will tend to score higher than those who would be indifferent or displeased.

Here we wished to investigate the differential effect on ESP scores of varying degrees of satisfaction at the idea of possessing ESP ability. Presumably, the individual who is more ego-involved in the demonstration of ESP will score higher than others. The following question was therefore asked:

What would be your attitude if your score on this test contributed substantial evidence of the existence of ESP?

- | | | |
|-------------------------------------|----|-------------------------|
| (a) I would be pleased | 60 | } (number of responses) |
| (b) I would be entirely indifferent | 13 | |
| (c) I would not be pleased | 1 | |

The breakdown is presented in Table 8. There was only one subject who gave a displeasure response, and therefore he was added to the second group. As can be seen, the question did not differentiate well, since about 75 per cent fell in the pleased group. The deviations are not significant, nor are the differences between them.

- d. Subjects who show confidence in their ESP ability will tend to score higher than those subjects who do not show confidence.

TABLE 8
EFFECT OF PLEASURE AT HIGH SCORE ON
ESP RESULTS (SERIES 3 AND 4)

High Score Would Cause —	Number of Subjects	Total Calls Registered	Deviation	Excess of Deviation over Expectation, in %	CR	p
Pleasure	60	17752	+47.6	1.34%	<1	—
Indifference or Displeasure	14	4162	+ 6.6	0.79%	<1	—

It will be recalled from Schmeidler's work that one of the ways in which she described sheep was "those who expected to succeed." Furthermore, the general attitude of expectation of above-chance scoring would seem to be one of the prerequisites of any of the favorable attitudes for ESP that the experimenters have clinically observed.

It was therefore decided to test the subjects' attitudes in this area. In the five-hundred-mile Series 2, the following question was asked of the subjects:

Do you personally expect to get more correct hits than you would expect to get from chance alone?

The subjects were instructed that they had to answer "Yes" or "No" even if they were actually undecided about it. It was felt that by forcing the subjects to commit themselves, a "task set" (attitude toward success in the task) would be in some sense established.

The results of Series 2 are presented in Table 9.

TABLE 9
EFFECT OF CONFIDENCE IN ESP ABILITY
ON ESP SCORE (SERIES 2)

Expect To Be Above Chance?	Number of Subjects	Total Calls Registered	Deviation	Excess of Deviation over Expectation, in %	CR	p
"Yes"	43	6416	+34.8	2.71%	1.09	.14
"No"	20	2983	+40.4	6.77%	1.80	.036

As can be seen from Table 9, the question did not yield the expected differentiation; in this case, the actual relationships were reversed, although not significantly so. The difference between the means of the groups does not approach significance.

On the basis of these results, it was felt that perhaps the undecided subjects should not be forced into one of the two alternative categories, since if this rather artificial choice failed to have any relation to ESP, their scores might be hiding the expected deviations in the case of those subjects who were willing to make a definite commitment, and whose attitudes, being definite, might with more justification be expected to affect their scoring levels. Therefore in Series 3 and 4, the following question was asked:

Approximately one out of every four subjects scores significantly above chance. How do you personally expect to score?

- A. (a) I am quite sure I will score above chance 4
 (b) I probably will score above chance 16
 (c) I probably won't score above chance 28
 (d) I am quite sure I won't score any better than
 chance 3
- B. If you are completely unable to make any estimate
 concerning your score, check 23

The statement that one in four subjects scored well was made not as a statement of actual fact, but merely to attempt to establish a common standard of reference in terms of which the subjects could rate themselves.

It can be seen from the distribution of answers that this question failed to differentiate well, since only seven subjects fell in the two extremes. For this reason, we considered the response in terms of a three point scale. The results are presented in Table 10.

TABLE 10
EFFECT OF CONFIDENCE IN ESP ABILITY
ON ESP SCORE (SERIES 3, 4)

Expect to Score —	Number of Subjects	Total Calls Registered	Deviation	Excess of Deviation over Expectation, in %	CR	p
Above Chance (A; a, b)	20	5917	+17.6	1.49%	<1	—
Don't Know (B)	23	6890	+37.0	2.69%	1.11	.13
At Chance (A; c, d)	31	9107	— 0.4	—	—	—

As can be seen from Table 10, none of the three groups scored significantly, nor was there a significant difference between any of them. This is consistent with the findings in Table 9, where similarly no significant differences were found.

- e. Expansive subjects will tend to score above chance on direct hits in ESP tests of the clairvoyance type, while compressive subjects will tend to score at or below chance.

This hypothesis is based on the work of Humphrey (8, 9, 10). It will be recalled that in all the published work concerning this hypothesis, there have been successful differentiations made when pictures have been used as the targets instead of cards. When cards have been used as targets, there have been experiments in which the subjects could not be differentiated on the basis of drawings made before the ESP tests.

To attempt to test this hypothesis, subjects in Series 3 and 4 were asked to try to reproduce a picture which had been placed in a manila envelope at the front of the room. They were merely told that it was a picture taken from a magazine. This was not meant to be an ESP test, but merely a situation which the subject would think was an ESP test. We were not interested in how the individual reacted to *any* social situation, but only to a specific type of situation—the ESP situation. It is quite possible that responses would differ in the two different situations. The pictures were rated on the basis of the criteria offered by Elkich in her monograph (2). They were rated by two individuals, working together, since it was felt that it would not be wise for two entirely inexperienced judges to attempt to rate them independently and then expect to achieve any high degree of reliability. The results are presented in Table 11.

TABLE 11
COMPARISON OF EXPANSIVE AND COMPRESSIVE
SUBJECTS (SERIES 3, 4)

Type of Subject	Number of Subjects*	Total Calls Registered	Deviation	Excess of Deviation over Expectation, in %	CR	p
Expansives	50	14780	+46.0	1.56%	<1	—
Compressives	23	6835	+ 6.0	.44%	<1	—

As can be seen from Table 11, while there is a difference in the expected direction between the two groups, it is not significant.

* One subject failed to make drawing and was therefore omitted from table.

It should be noted that Humphrey has never attempted to rate her subjects on the basis of a single picture; at the very least she has used two. This may account for the poor results here. It should be further realized that perhaps the inexperience of the judges favored poor judgments as to the proper ratings. For these reasons our results are not comparable with Humphrey's work.

- f. Subjects who report they would be disturbed by the possession of ESP ability will tend to score below chance expectation, while those who report they would not be disturbed will tend to score above chance expectation.

While this would seem to some to be the most direct type of approach to the question of which individuals will demonstrate ESP, it has not been used in any published studies. It should be realized that personality ratings based upon general situations may or may not be found to be valid for the subject when in the ESP situation. For example, a deeply introspective, imaginative, withdrawn person might be found to be very poorly adjusted as indexed by the Rorschach test. Nevertheless, he might feel very much at home in the ESP situation, and the thought of extending his perceptive capacities might be a very intriguing and satisfying one to him. On the other hand, an outgoing, socio-centered, extroverted type of individual might be very well adjusted as indexed by the Rorschach, but when confronted by the ESP situation feel quite insecure in being asked to attempt a task so foreign to his normal activities.

An analogous point was recently made by Schmeidler in a paper correlating ESP scores and frustration-aggression ratings on the *Rosenzweig Frustration Test* (32). She pointed out that the aggression rating must be considered in light of the testing situation itself, as well as of any general aggressive trait the test might be expected to measure.

One of the most fundamental hypotheses in clinical psychology concerns the defensive behavior characteristic of an individual who feels threat, apprehension, or anxiety. Therefore, there is good reason, if we assume that ESP expresses the dynamics of personality, to look for the operation of this pattern in extrasensory perception. We would expect that if a subject perceived "reaching out for" or "opening himself up to" paranormal impressions as carrying with it some sort of vague or specific threat, then he would in consequence refuse defensively, perhaps unconsciously, to use extrasensory perception. On the other hand, the individual who does not feel apprehensive over the possibility of possessing ESP would be expected to tend to demonstrate ESP, all other things being equal.

A rough attempt to do this was made by asking the following question:

What would be your attitude if your score on this test contributed substantial evidence of the existence of ESP?

- (a) I would be somewhat disturbed 9
 (b) I would not be disturbed 62

The breakdown is presented in Table 12.

TABLE 12

EFFECT OF EMOTIONAL REACTION TO SCORING WELL
ON ESP SCORE (SERIES 3, 4)

High Score Would Make Subject —	Number of Subjects*	Total Calls Registered	Deviation	Excess of Deviation over Expectation, in %	CR	p
Disturbed	9	2694	-28.8	-5.35%	1.40	.08
Not Disturbed	62	18325	+77.0	2.10%	1.42	.08
<i>Difference</i>						
$CR_{diff} = 1.80$						
$p_{diff} = .036$						

From Table 12 we see that the average deviations of each group were in the predicted direction, and that the CR of the difference between the means of the groups was significant at the 3 per cent level of confidence. We may view these results as suggestive.

The writer should like to use this as a specific example of the type of thing one might expect to find if modern personality theory were systematically applied in parapsychology. We might expect to find subjects who responded to the ESP task with apprehension or anxiety rejecting or displacing the targets, regardless of whether they were rated as *generally* adjusted or not. In fact, it should be noted that any observed correlation between general adjustment and ESP score is in all probability a byproduct of two other correlations: between scoring level and adjustment to the specific situation on the one hand, and between this specific adjustment and general adjustment on the other.

It should be constantly remembered that there is no *empirical reason* to believe that ESP expresses the dynamics of personality. Murphy,

* Three subjects gave no response to this question and were therefore omitted from table.

for example, has merely expressed the opinion that if we *assume* that extrasensory perception is similar to sensory perception in terms of its relation to personality structure, predictions can be made as to where and how to go about looking for the factors which control ESP. *Prima facie* it would seem a reasonable assumption, and therefore one which most experimenters have been willing to make. So if we make this assumption, why not make full use of all it implies? It seems appropriate to take the numerous hypotheses concerning sensory perception which have been worked out by dynamic psychology, and extensively apply them to extrasensory perception. This is the sort of thing we have tried to illustrate by bringing to bear on ESP our knowledge of the relation of anxiety to defensive behavior.

It should not be forgotten, however, that it is quite possible that we shall ultimately discover that personality plays an extremely minor role in ESP, and that actually we are dealing with a physiological threshold effect of some sort, only slightly amenable to any autistic processes. Therefore, it should be emphasized that all parapsychological research being carried on against the background of dynamic psychology is purely exploratory; for the present question is not, "*How* does ESP express the dynamics of personality?" but merely, "*Does* ESP express the dynamics of personality?" If the answer to the latter question is finally, "Yes, but not enough to make a point of it," then the physiologist may ultimately find that he has inherited the psychologist's problem. Our present work is largely based on an argument by analogy, and we cannot treat our personal apperceptions of the nature of ESP as certainties until we are empirically justified in doing so.

Conclusions

An attempt was made to see whether the attitudes a subject brings with him to an ESP test can be correlated with his level of scoring. Six attitudes were investigated, each attitude being registered on scales with from two to five points. The four and five point scales were found to differentiate poorly in terms of the response distribution, and were, therefore, tabulated in terms of a two or three point scale. In no case did the categories yield statistically significant differences in scoring.

Two of the attitudes tested had been reported in the literature as having successfully been used for purposes of differentiation by other workers. The first, the sheep-goat hypothesis, was found not to be stated in a form which could be clearly repeated. The second, the compressive-expansive hypothesis, had been developed on the basis

of ESP tests using drawings; when used in ESP tests using cards it had not always been successful. Furthermore, the drawing index used here was not as extensive as that used in the original experiments, and therefore introduced the possibility of more error.

The selection of three other attitudes was based largely on the clinical impressions which abound in the literature. The fourth was selected in an attempt to extend a specific hypothesis from normal psychology to the case of ESP.

The predominantly negative conclusions which we have reached could be criticized on several grounds. It might be said that there is no evidence that ESP was definitely operating in Series 3 and 4 (see Table 1), and therefore the negative results are meaningless. This is an important point, but it must be kept in mind that the studies from which several of the hypotheses were formulated did not give significant total deviations either; i.e., experiments which were insignificant in terms of total ESP were used to demonstrate that ESP was actually operating in a more subtle manner by using the differentiating technique discussed above. It does not seem justified to accept an insignificant experiment if one breakdown of the data supports a hypothesis, but to reject such an experiment as invalid if the same breakdown does not support the hypothesis.

A second criticism is also important. It might be said that actually the hypotheses tested are valid, but that the methods used to verify these hypotheses were inadequate or invalid. This is a quite legitimate point to make; however, attitudes can only be talked about if they can be operationally defined, for otherwise the concept of attitudes is useless. Therefore, a criticism of these indexing methods, which are admittedly crude, should contain a suggestion as to better indexing methods. It should be noted that it is mere circularity to define favorable attitudes as those which are present when ESP is demonstrated. For the word "attitudes" to have any nontautological meaning, they must be defined independently of ESP. This, of course, has not always been done in the literature, where there has been the occasional tendency to hypothesize the presence of the "wrong" attitudes or motivational systems when a subject or group failed to demonstrate ESP.

The fact that most of the deviations were in the predicted directions can be interpreted in the following way: that if real variables are operating to give these differences, they exert such a weak effect that it is very doubtful that they will ever be of any practical use to parapsychology.

The most important conclusion that we can draw from the failure of this hypothesis to be clearly verified is the following. The wide-

spread opinions held by parapsychologists as to the extreme importance of proper attitudes in the demonstration of ESP should be looked upon with some suspicion, and should, therefore, be reconsidered by extensive empirical investigations.

It may well prove to be a case in which opinions must be eventually revised, or even in some instances reversed, when they are considered in the light of experimental investigations. This investigation should be a top priority research goal, for so much is made of attitudes in the literature as the basis of the unrepeatability of ESP experiments, that, if this emphasis has been unjustified, we should immediately become aware of it in order to concentrate upon other areas.

Hypothesis VI

Subjects who are rated as well adjusted will tend to score above chance, while subjects who are rated as poorly adjusted will tend to score below chance.

Methods and Results

One of the better validated hypotheses in parapsychology concerns the correlation between personality adjustment and ESP scores. Humphrey (7) found suggestive correlations using the Bernreuter Inventory, and Schmeidler (28, 29, 31) found highly significant correlations between ESP and Rorschach scores using Munroe's "inspection" technique. Therefore, it was felt that it would be desirable if similar correlations could be found independently. Unfortunately, it was impractical to make use of the Rorschach test. Therefore, it was decided to use a personality inventory of some sort. Upon the advice of Dr. Dyer of the Harvard Office of Testing, it was decided to make use of the recently devised Heston Personal Adjustment Inventory (5). This inventory consists of 270 questions. These questions are meant to index six basic components of an individual's adjustment. It has advantages over a test such as the Bernreuter in that no question is used on more than one scale; thus, correlations between any of the six scales are not due to statistical artifacts. The six scales have reliability coefficients ranging between .8 and .9. The scales were validated by means of the method of internal consistency, the psychological meaningfulness of the component items, and by comparisons with independent criteria derived from clinical judgments. The latter correlations seem high enough to warrant some faith in the Inventory.

The six traits are named and described briefly as (1) Analytic thinking (thinking introversion); (2) Sociability (social extraversion); (3) Emotional stability; (4) Confidence; (5) Personal relations (indicating congeniality and non-irritation at the behavior of others); (6) Home satisfaction.

Low scores in these areas imply poor adjustment. The scoring is based on norms developed for college populations. It is given in percentiles. Several of the scales are highly correlated with one another. Heston justifies the use of the several scales in these cases of high correlation instead of a single scale on the grounds that the scales are derived independently, and moreover, that they are not correlated so highly as to be useless individually when a particular subject's adjustment is being tested. Since the scales were so highly correlated, and since there was no valid way of predicting separations

in terms of individual scales on the basis of the Humphrey or Schmeidler work, it was decided that a single criterion would be used in prediction. This was individual adjustment, derived from averaging the decile scores for each of the six trait scales. It was felt that this would be the closest method of approaching some degree of similarity to the adjustment hypotheses in the literature. It was further felt that this would help reduce the error inherent in personality inventories. It was predicted on the basis of Humphrey's work with a personality index that poorly adjusted subjects would score definitely below chance rather than fluctuate around chance, and that well adjusted subjects would score above chance, as reported by Schmeidler using the Rorschach. It was felt that this study would be more similar to Humphrey's study than to Schmeidler's.

The results are presented below. The five subjects who scored exactly at chance expectation were omitted from the four-way table; the subjects whose scale average placed them in the fifth decile or below were defined as poorly adjusted; the subjects whose scale average placed them above the fifth decile were defined as well adjusted. It was realized that this was a crude method of indexing the subjects. Nevertheless, it was hoped that if personality differences were correlated with ESP scores they would show up by using the indexing techniques described.

The results are presented below in Table 13. The chi-square method was used to estimate significance.

TABLE 13
EFFECT OF PERSONAL ADJUSTMENT
ON ESP SCORE (SERIES 4)

Personal Adjustment	ESP Score Below Chance	ESP Score Above Chance	Total No. of Subjects*
Below Average (lower five deciles of Heston Scale)	19	8	27
Above Average (upper five deciles of Heston Scale)	7	13	20
Total Number of Subjects*	26	21	47

*5 subjects who scored exactly at chance are omitted from the table.

$$\chi^2 = 4.47 \quad (\text{with Yates small-n correction})$$

$$p = .038 \quad (\text{Degrees of freedom, 1})$$

Conclusions

As can be seen from Table 13, the subjects who scored above chance and the subjects who scored below chance could be differentiated to some degree by use of a personality inventory designed to measure personal adjustment. However, the difference does not reach the .01 level of significance. To the extent that it can be considered significant, it substantiates the hypothesis.

General Summary

A number of research studies were reported. These studies attempted to do several things:

1. To construct a testing technique which adequately met the usual counter-hypotheses to ESP.
2. To choose a number of well-verified hypotheses from the literature, and to attempt to repeat the relevant earlier procedures.
3. To attempt to demonstrate several new experimental approaches to ESP.

The primary findings were encouraging; it was concluded that extrasensory perception had been observed to operate under conditions in which no counter-hypothesis could be offered to invalidate the conclusions. Further, several of the best verified hypotheses were again verified under these ideal experimental conditions.

Only in the case of one series of hypotheses, which state that attitudes on the part of the subject are correlated with scoring level, were non-significant results obtained. These results, in the case of the two subhypotheses in this area which were based upon experimental findings, were understood in one case in terms of the possibility that the testing methods used here were inadequate, and in the other case, in terms of the fact that the hypothesis was so stated in the literature as to be currently difficult of precise test. In the case of the hypotheses which were based on clinical impressions, the failure to achieve positive results were interpreted to suggest that many heretofore accepted opinions of many parapsychologists as to the relevance of attitude and motivation to extrasensory perception should be looked into further. This implies that not just the particular subhypotheses tested in this area are in question, but that the validity of the original hypothesis—the relevance of attitudes in general, whether toward ESP itself, or toward the experimenter, or toward the testing situation—is under question. Therefore, it was concluded that the whole area concerned here should be submitted to a detailed experimental investigation.

However, aside from this one negative finding, this experimental study can be justifiably considered to support all the hypotheses of experimental origin which it was adequately able to investigate.

This report would seem to be in the nature of a convincing verification of these basic ESP hypotheses. Further, since it was carried on quite independently of other investigators and their methods, the conclusion can be drawn that these hypotheses were not only verified, but independently verified.

Thus, these results can be considered as extremely encouraging in terms of the attempts of parapsychologists in the last fifteen years to put the demonstration of ESP on a repeatable basis.

Appendix I

A Study of Decline Effects in ESP

S DAVID KAHN AND ULRIC NEISSER

Introduction

The following is a preliminary report of a large-scale investigation of the relationship between deep-level personality variables and ESP. This relationship is considered by many parapsychologists to be one of the most important research problems in the field. The personality data were collected as part of a systematic and extended research program into personality structure carried out by a department of Harvard University. When these data have been collated, new material will be available which may shed further light upon the important ESP-personality relationship.

Most of the ESP work was done with the IBM scoring technique. It was felt advisable to report these results at the present time because of their bearing on the earlier IBM work already discussed in these *Proceedings*, although the main experimental material has not yet been analyzed.

In addition to the personality research, we were particularly interested in studying the effects of repeated ESP testing upon performance and the decline in scoring which might be expected to occur in successive tests with the same subjects. Extensive decline effects within the single IBM sheet, in terms of group averages, have already been reported above (in Table 2 of these *Proceedings*). In view of the long-term declines in card experiments that have been found by several experimenters (30) it seemed reasonable to expect a systematic decline to occur when subjects were submitted to a lengthy series of IBM tests.

Accordingly, the experimental program required our nineteen subjects to fill out six IBM sheets, during six different weeks, a new target being presented each week.

Procedure

Each of the subjects was first interviewed individually for one hour. At that time, the interviewer ascertained informally the subject's attitude toward ESP, and inquired as to any spontaneous experiences the subject could remember which seemed to bear upon paranormal phenomena. The subject was then given a Stuart-type picture test (36) in which he was asked to do his own matching.

At the conclusion of the interview, he was given an IBM sheet to take with him. He was informed that the target IBM sheet was in a locked metal box in a specified place—this box was shown him—and would remain there for the following week. He was instructed to attempt to duplicate this target at some convenient time during that week. He was told to work at whatever rate and with whatever "system" appealed to him. It was emphasized that he should choose a time when he was relaxed and would not be disturbed, and a time when he "felt like doing it." He was asked to return the completed sheet at the end of the week. He was informed that he would receive a new sheet in the mail the following Monday morning, and the whole procedure was to be repeated with a new target in place of the old. This was to be done for six weeks, consecutive except for a week's interruption for the spring vacation. During the entire interview, the interviewer refrained from making any comments indicating his own attitude on the question of ESP.

This plan was carried out successfully with seventeen of the subjects. Two of the subjects filled out only five sheets, missing one week.

The subjects were paid at the rate of one dollar for the hour's interview and fifty cents per IBM sheet. They were being paid at a similar rate for their participation in the personality tests.

This general procedure was followed because it seemed desirable to eliminate the experimenter as far as possible from the physical and psychological field of the subject during the actual test; for, although it is known that the experimenter has a definite effect on the performance of the subject when the two are in the relatively close relationship which exists in the ordinary ESP experiment, we have as yet no means of controlling or measuring this effect. It was hoped that the "remote control" method would at least partially eliminate the effect of this variable. In addition, this plan made possible a degree of spontaneity and relaxation in the subject nearly unobtainable in a laboratory test held at an appointed time.

No results were given the subjects during the entire course of the experiment.

Results

Table 1, below, contains the total results for all nineteen subjects, summed for the separate weeks and for the two sides of the answer sheets. These data are graphed in Figure 1. It may be observed that both of the expected types of decline effects actually occur. In each of the six different weeks, scoring is higher on the first halves of the answer sheets than on the second. The first-half as well as the

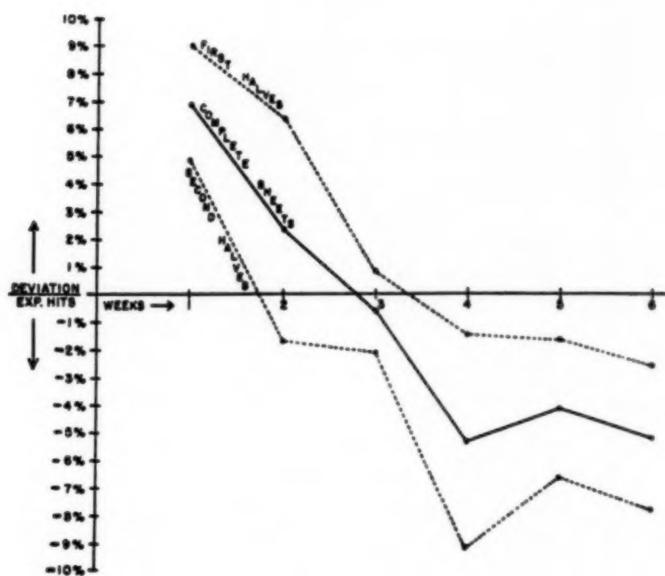


FIGURE 1

Total results, for all weeks, and for the two halves of the sheets.

second-half totals decline steadily from week to week as the experiment progresses, with one exception: a rise in the second-half total in the fifth week.

The difference between the positive deviation of 64.2 for all the first-halves and the negative deviation of 113.8 for the second-halves is significant at the .007 level.

As seen in Table 2 the difference between the positive deviation of 107.4 on the first three weeks combined and the negative deviation of 157.0 on the last three weeks is significant at well beyond the .001 level. Thus there can be no doubt that the declines observed are due to the operation of extrasensory perception in this experiment.

As analyzed in Tables 1 and 2, these declines emerge from the group totals. In Tables 3 and 4 the declines across the two sides of the sheet and the two parts of the series are presented for the individual subjects. Of the nineteen subjects, fifteen were higher in the first three weeks than in the last three weeks. Fourteen were higher on the first halves of their sheets than on the second halves of their sheets. Thus, the majority of individual subjects duplicated the group

TABLE 1

DECLINE EFFECTS FROM SIDE 1 TO SIDE 2

Week	Total Calls Registered	Deviation	Excess of Deviation over Expectation, in %	CR	p
<i>Side 1</i>					
<i>Calls 1-150</i>					
1.	2845	+ 50.0	8.79%	2.34	.01
2.	2848	+ 36.4	6.39%	1.71	.04
3.	2690	+ 6.0	1.12%	<1	—
4.	2844	— 6.8	—1.20%	<1	—
5.	2694	— 6.8	—1.26%	<1	—
6.	2848	— 14.6	—2.56%	<1	—
Totals	16769	+ 64.2	1.91%	1.24	.11
<i>Side 2</i>					
<i>Calls 151-300</i>					
1.	2841	+ 27.8	4.89%	1.30	.09
2.	2813	— 3.6	—0.64%	<1	—
3.	2691	— 9.2	—1.71%	<1	—
4.	2843	— 50.6	—8.90%	2.37	.009
5.	2694	— 33.8	—6.27%	1.63	.05
6.	2847	— 44.4	—7.80%	2.08	.02
Totals	16729	—113.8	—3.40%	2.20	.012
Grand Totals	33498	— 49.6	—0.74%	<1	
<i>Side 1</i>					
CR = 1.24					
p = .11					
<i>Side 2</i>					
CR = 2.20					
p = .012					
<i>Difference</i>					
CR _{diff} = 2.46					
p _{diff} = .007					

TABLE 2
SERIAL DECLINE EFFECTS

Week	Total Calls Registered	Deviation	Excess of Deviation over Expectation, in %	CR	p
<i>First Three Weeks</i>					
1.	5686	+ 77.8	6.84%	2.58	.005
2.	5661	+ 32.8	2.90%	1.09	.14
3.	5381	- 3.2	-0.30%	<1	—
Totals	16728	+107.4	3.21%	2.07	.02
<i>Second Three Weeks</i>					
4.	5687	- 57.4	-5.05%	1.90	.03
5.	5388	- 40.6	-3.77%	1.35	.09
6.	5695	- 59.0	-5.18%	1.95	.03
Totals	16770	-157.0	-4.68%	3.03	.001
Grand Totals	33498	- 49.6	-0.74%	<1	—
<i>First Three Weeks</i>					
CR = 2.07					
p = .02					
<i>Second Three Weeks</i>					
CR = 3.03					
p = .001					
<i>Difference</i>					
CR _{diff} = 3.64					
p _{diff} = .00014					

TABLE 3
DECLINE EFFECTS FROM SIDE 1 TO SIDE 2
BY INDIVIDUAL SUBJECTS

Subject	Side 1: Calls 1-150			Side 2: Calls 151-300		
	Total Calls Registered	Deviation	%	Total Calls Registered	Deviation	%
Beep	900	+ 5.0	2.8%	899	— 0.8	— 4.5%
Bruise	897	+13.6	7.6%	900	— 5.0	— 2.8%
Cracker	894	+ 7.2	4.0%	866	— 2.2	— 1.3%
Chalk	900	+11.0	6.1%	897	— 6.4	— 3.6%
Chug	898	+ 4.4	2.5%	898	— 10.6	— 5.9%
Choice	899	+ 5.2	2.9%	895	— 4.0	— 2.2%
Grope	895	— 4.0	— 2.2%	894	— 23.8	—13.3%
Heiner	899	+ 2.2	1.2%	899	+ 3.2	1.8%
Halt	900	+26.0	14.4%	900	— 18.0	—10.0%
Hudge	745	+ 6.0	4.0%	743	— 8.6	— 5.8%
Jeave	898	—16.6	— 9.2%	900	— 2.0	— 1.1%
Nob	900	— 3.0	— 1.7%	900	+ 14.0	7.8%
Neater	750	—11.0	— 7.7%	749	— 23.8	—15.9%
Stall	899	+31.2	17.4%	898	— 10.6	— 5.9%
Tally	898	+ 8.4	4.7%	898	— 0.6	— 0.3%
Toast	899	—27.8	—15.5%	900	— 2.0	— 1.1%
Thaw	899	+ 6.2	4.0%	897	— 12.4	— 6.9%
Yeast	899	+ 7.2	4.0%	898	+ 0.4	0.2%
Zoid	900	— 7.0	— 3.9%	898	— 0.6	— 0.3%
Totals	16769	+64.2	1.9%	16729	—113.8	— 3.4%

	Number of subjects whose scores are higher on side 1	Number of subjects whose scores are higher on side 2
Actual numbers	14	5
Chance expectancy	9.5	9.5
$\chi^2 = 4.26$ (degrees of freedom, 1)		
$p = .04$		

TABLE 4

SERIAL DECLINE EFFECTS BY INDIVIDUAL SUBJECTS

Subject	First Three Weeks			Second Three Weeks		
	Total Calls Registered	Deviation	%	Total Calls Registered	Deviation	%
Beep	899	+ 15.2	8.5%	900	- 11.0	- 6.1%
Bruise	900	- 5.0	- 2.8%	897	+ 13.6	7.6%
Cracker	864	+ 10.2	5.9%	896	- 5.2	- 2.9%
Chalk	900	- 1.0	- 0.6%	897	+ 5.6	3.1%
Chug	898	+ 8.4	4.7%	898	- 14.6	- 8.1%
Choice	896	+ 3.8	2.1%	898	- 1.6	- 0.9%
Grope	891	- 7.2	- 4.0%	898	- 20.6	-11.5%
Heiner	899	+ 20.2	11.2%	899	- 14.8	- 8.2%
Halt	900	- 11.0	- 6.1%	900	+ 19.0	10.6%
Hudge	895	+ 6.0	3.4%	593	- 8.6	- 7.3%
Jeave	899	- 0.8	- 0.4%	899	- 17.8	- 9.9%
Nob	900	+ 17.0	9.4%	900	- 6.0	- 3.3%
Neater	600	- 1.0	- 0.8%	899	- 33.8	-18.8%
Stall	898	+ 7.4	4.1%	899	+ 13.2	7.3%
Tally	899	+ 22.2	12.3%	897	- 14.4	- 8.1%
Toast	899	- 2.8	- 1.6%	900	- 27.0	-15.0%
Thaw	896	+ 7.8	4.4%	900	- 14.0	- 7.8%
Yeast	897	+ 18.6	10.4%	900	- 11.0	- 6.1%
Zoid	898	- 0.6	- 0.3%	900	- 8.0	- 4.4%
Totals	16728	+107.4	3.2%	16770	-157.0	- 4.7%

	Number of subjects whose scores are higher in the first 3 weeks	Number of subjects whose scores are higher in the second 3 weeks
Actual numbers	15	4
Chance expectancy	9.5	9.5
$\chi^2 = 6.37$ (degrees of freedom, 1)		
$p = .012$		

trends. In Tables 3 and 4 the chi-square analysis is applied to the data in terms of the number of subjects showing predicted declines. It may be seen that the serial decline effect is significant here at the .02 level; the decline from the first half to the second half of the Record Sheet is at the .04 level. Thus evidence for the reality of the decline effects is provided not only by the group totals but by the responses of the subjects considered individually.¹

Conclusions

A preliminary report has been presented of an experiment studying the relationship between ESP and personality, as well as certain decline effects in ESP itself. Only the latter, as demonstrated by a series of six IBM tests with the same group of nineteen subjects are reported here. It is found that, under conditions which exclude sensory leakage or scoring error, two strong types of decline occur in the degree of coincidence between subject's record and target. In the first place, the second side of the answer sheet tends to score lower than the first. In the second place, there is a rather steady decline in scoring as the experiment progresses, so that the scores of the first three weeks are considerably higher than those of the second. These two effects are highly significant in the group taken as a whole. They are significant at a lower confidence level if individual subjects are considered in a chi-square analysis. There can be no doubt that ESP was operating in this experiment. Decline effects of this sort may be considered typical of the operation of ESP under these conditions.

¹ The figures presented here are not the direct results of the original machine check. The data were rechecked by Mrs. L. A. Dale, Research Associate of the Society, to make a hand check available for comparison with the machine. There were a number of discrepancies, which were carefully reexamined to determine the accurate totals for presentation here.

The discrepancies between the original machine check and the final accurate totals were: (a) 10 errors of omission and commission not noted originally, and (b) 21 errors in the number of hits, of which ten occurred with subject "Jeave" and five with subject "Tally." These latter fifteen were all underestimations by the machine. A specialist in the IBM Test Scoring Machine was consulted, and suggested that these fifteen were due to improper marking of the sheets by the subjects. Adequate preparation of the sheets for the machine, as indicated in the main section of the present paper, would have prevented them, and would also have assured accuracy in the count of omissions and commissions. Hence the errors inherent in the IBM technique itself total 6, out of 6650 hits, a proportion (.09%) not unlike that found in the data of the main section. This may be taken as further evidence of the accuracy of the machine scoring method, when properly applied.

The differences between the original check and the final accurate totals were not significant. The CR_{diff} for the decline from side 1 to side 2 changed from 2.57 to a final value of 2.46. The CR_{diff} for the serial decline effect was raised from 3.63 to 3.64 by the corrections. The chi-squares in Table 3 and 4 were unaltered.

Appendix III

HARVARD SOCIETY FOR PARAPSYCHOLOGY

ATTITUDE QUESTIONNAIRE

Name: Age: Sex:

Address:

Please be sure to answer all the questions.

1. Do you think that extrasensory perception (ESP) is theoretically possible:

I. in this particular experiment? CHECK ONE.

Yes

No

II. under other circumstances: CHECK ONE.

Yes

No

2. Do you think it probable that extrasensory perception ever actually occurs? CHECK ONE.

(a) Very probable

(b) Fairly probable

(c) Slightly probable

(d) Improbable

(e) Very improbable

3. What would be your attitude if your score on this test contributed substantial evidence of the existence of ESP?

I. CHECK ONE.

(a) I would be pleased

(b) I would be entirely indifferent

(c) I would *not* be pleased

II. CHECK ONE.

(a) I would be somewhat disturbed

(b) I would not be disturbed

4. Approximately one out of every four subjects scores significantly above chance. How do you personally expect to score? CHECK ONE.

A.

(a) I am quite sure I will score above chance

(b) I probably will score above chance

(c) I probably won't score above chance

(d) I am quite sure I won't score any better than chance

B.

If you are completely unable to make any estimate concerning your score, check:

5. Are the opinions registered above based on previous experiences with ESP? If so, please describe them briefly on the other side of this sheet.

6. Do you feel that you have ever had any psychic experiences? If so, please describe briefly.

7. Have there ever been any in your family? If so, please give a brief description.

THANK YOU.

REFERENCES

1. Bevan, J. M. The relation of attitude to success in ESP scoring. *J. Parapsychol.*, 1947, 11, 296-309.
2. Elksch, P. Children's drawings in a projective technique. *Psychol. Monog.*, 1945, 58, No. 1.
3. Grela, J. J. Effect on ESP scoring of hypnotically induced attitudes. *J. Parapsychol.*, 1945, 9, 194-202.
4. Greville, T. N. E. On multiple matching with one variable deck. *Ann. Math. Statist.*, 1944, 15, 432-434.
5. Heston, J. C. *Heston Manual and Personal Adjustment Inventory*. World Book Co., Yonkers-on-Hudson, 1949.
6. Humphrey, B. M. Patterns of success in an ESP experiment. *J. Parapsychol.*, 1943, 7, 5-19.
7. ———. Introversion-Extraversion ratings in relation to scores in ESP tests. *J. Parapsychol.*, 1951, 15, 252-262.
8. ———. Success in ESP as related to form of response drawings: I. Clairvoyance experiments. *J. Parapsychol.*, 1946, 10, 78-106.
9. ———. Success in ESP as related to form of response drawings: II. GESP experiments. *J. Parapsychol.*, 1946, 10, 181-196.
10. ———. The relation of ESP to mode of drawing. *J. Parapsychol.*, 1949, 13, 31-47.
11. Kahn, S. D. and Neisser, U. A mechanical scoring technique for testing GESP. *J. Parapsychol.*, 1949, 13, 177-185.
12. MacFarland, J. D. Discrimination shown between experimenters by subjects. *J. Parapsychol.*, 1938, 2, 160-170.
13. Munroe, R. L. Prediction of the adjustment and academic performance of college students by a modification of the Rorschach method. *Applied Psychol. Monog.*, 1945, No. 7.
14. "Random Number Tables." Prepared by the Harvard Psychological Laboratory, from Statistical Tables. Fisher and Yeats. Oliver and Boyd, London, 1943, Table 33.
15. Rhine, J. B. The effect of distance in ESP tests. *J. Parapsychol.*, 1937, 1, 172-184.
16. ———. *Extra-Sensory Perception*. Boston Society for Psychic Research, 1934, 21-22.
17. ——— and associates. *Extra-Sensory Perception After Sixty Years*. Henry Holt and Co., New York, 1940, 304-305.
18. *Ibid.* 266-290.
19. *Ibid.* 338-341.
20. *Ibid.* 24.
21. *Ibid.* 106-180.
22. *Ibid.* 182-212.
23. Scherer, W. B. Spontaneity as a factor in ESP. *J. Parapsychol.*, 1948, 12, 126-147.
24. Schmeidler, G. R. Predicting good and bad scores in a clairvoyance experiment: A preliminary report. *J. A.S.P.R.*, 1943, 37, 103-110.
25. ———. Predicting good and bad scores in a clairvoyance experiment: A final report. *J. A.S.P.R.*, 1943, 37, 210-221.
26. ———. Separating the sheep from the goats. *J. A.S.P.R.*, 1945, 39, 47-49.
27. ———. Progress report on further sheep-goat series. *J. A.S.P.R.*, 1946, 40, 34-35.
28. ———. Rorschach variables in relation to ESP scores. *J. A.S.P.R.*, 1947, 41, 35-64.

48] *Proceedings of the American Society for Psychical Research*

29. ———. Personality correlates of ESP as shown by Rorschach studies. *J. Parapsychol.*, 1949, 13, 23-31.
30. ———. Position effects as psychological phenomena. *J. Parapsychol.*, 1944, 8, 110-123.
31. ———. Comparison of ESP scores with Rorschachs scored by different workers. *J. A.S.P.R.*, 1949, 43, 94-97.
32. ———. Some relations between picture-frustration ratings and ESP scores. *J. Personal.*, 1950, 18, 331-344.
33. ——— and Murphy, G. The influence of belief and disbelief in ESP upon individual scoring levels. *J. Exp. Psychol.*, 1946, 36, 271-276.
34. Smith, B. M. and Humphrey, B. M. Some personality characteristics related to ESP performance. *J. Parapsychol.*, 1946, 10, 269-289.
35. Stuart, C. E. An analysis to determine a test predictive of extra-chance scoring in card-calling tests. *J. Parapsychol.*, 1941, 5, 99-137.
36. ———. An ESP test with drawings. *J. Parapsychol.*, 1942, 6, 20-43.
37. ———, Humphrey, B. M., Smith, B. M. and McMahan, E. Personality measurements and ESP Tests with cards and drawings. *J. Parapsychol.*, 1947, 11, 118-146.
38. Taves, E., and Dale, L. A. The Midas touch in psychical research. *J. A.S.P.R.*, 1943, 37, 57-83.
39. Woodruff, J. L., and Dale, L. A. Subject and experimenter attitudes in relation to ESP scoring. *J. A.S.P.R.*, 1950, 44, 87-112.

