THE QUESTION OF SENSORY CUES AND THE EVIDENCE

J. B. RHINE

ABSTRACT: Experimental conditions are chosen to exclude definitely all possible sensory cues: In clairvoyance tests with packs of 5-suit cards the subjects are not to see or touch the cards during the entire run. All the results from tests made under this narrowly defined set of conditions are then surveyed, in three general groups: I, opaque envelopes; II, screened untouched cards; III, distance tests with walls between. In the 13 series found to meet the conditions the total deviation is large enough to be very significant. The odds are 10^{196} to one that so high a score would not result from chance. Both the hypothesis of sensory cues and that of chance, then, are inapplicable to the explanation of work done under these conditions.

THERE has recently been some little discussion as to whether the results of tests of ESP ability reported are significantly different from what would be expected by chance alone. It is a question only of the sufficient soundness and applicability of the methods of evaluation used. The indications are rather strong that there will shortly be quite general agreement that the hypothesis of chance is inadequate to explain these results.

If the results of tests for extra-sensory perception are not explainable by chance, it is proper to ask next whether the senses have been adequately excluded by the experimental conditions from influencing the scores. Those who have followed closely the published reports are familiar with the now extensive literature in which various advanced conditions in the tests are reported which appear to render sensory perception impossible. Some of the most outspoken critics commenting on the research seem, however, generally not to have considered the better test conditions, and it may be of advantage to assemble the data for the focus of their attention.

Still better justification for this survey is the evident need to take stock, for the benefit of those engaged in the research as well as those otherwise interested in it, of the most conclusive evidence against the hypothesis that the results published are explainable by sensory cues.

Such a treatment may well serve both as a target for helpful criticism and as a register of the state of progress made.

In order to deal effectively and with reasonable brevity with the question of sensory cues it is necessary, first, to limit the discussion to that hypothesis, making all other considerations secondary. Second, it must be clear at the start just what criteria of adequately safeguarding conditions will be followed in assembling the results of the experiments which bear upon the hypothesis.

If the criteria are made high enough and there are still significant results eligible, agreement on the exclusion of sensory cues is practically assured. To lift the question, then, to a level where agreement will be general it is necessary to use clear-cut standards of evidentiality which leave no ground for argument. In choosing the conditions which may be represented as barring out sensory perception, the error, if any, must be on the side of over-caution, even though large masses of valuable data will thereby be dropped from consideration on this immediate issue.[1] The lines drawn will seem unnecessarily exclusive and perhaps highly arbitrary to many experimenters and critics. For naturally all experimenters have been convinced that in their test conditions sensory cues were excluded. As a matter of fact, there has been no demonstrated grounds for doubting that this has been the case. But the question at this point is not, "How much data is there that may be reasonably regarded as extra-sensory?" but rather, "Is there *any* series of tests in which even the most drastic critic can see that the senses could not possibly have been factors?"

The criteria adopted in this survey for the classification of the evidence bearing on the question of sensory cues are as follows:

1. Only the tests of the extra-sensory perception of objects (clairvoyance) will be considered. All pure telepathy and GESP tests (in which both telepathy and clairvoyance are possible) will be rejected, no matter how well the screening may appear to have been or how far the distance between the subjects. This extreme discrimination is made partly to focus the issue and simplify judgment of the case for and against sensory cues. It is done in great part, too, because otherwise there would be difficulty and possible dispute as to what distance is adequate to exclude sensory cues of an auditory nature.

At the same time it leaves out not only the contributions of Estabrooks and those of Brugmanns in which sender and receiver were in different rooms with closed doors, but the work of Sinclair involving

[1] Valuable as such data may be on other questions concerning ESP. This theme will be resumed in a later paragraph.

thirty miles between subjects, that of Usher and Burt at still greater distances, and a considerable range of Duke work in telepathy and GESP conducted with short distances (subjects two rooms apart), the Junaluska series at 250 miles, and considerable work at still greater distances. This wholesale rejection of all tests permitting telepathy is done with full recognition that much of this work could be included without raising any question of sensory cues, but it would be at the expense of simplicity and clarity on the particular problem being now considered.[2]

2. Test results will be rejected unless the report shows that there was no vision or touching of the cards by the subject between the preparatory cut (following shuffling) by the experimenter and the end of the run of at least 25 calls or trials. The senses of vision and touch are those against which barriers are needed in clairvoyance tests. Since no one knows the card order (if cards are the basis of the tests), auditory cues are not possible.

Thereby is discarded again a huge block of data for which some investigators will properly maintain a stout defense: For instance, the DT tests in which the back of only the top card and the edges of the others are visible (unless careless handling exposes the bottom card to view), but in which the evidence shows that the hits significantly above the mean chance expectation are not confined to the top and bottom cards, as would be expected were sensory cues the explanation. Even those screened tests are left out in which the subjects handle the cards under or behind an opaque screen even though the subject only takes the card by thumb and forefinger by its corner or edge. Also excluded are the screened open matching performances by children reported by L. E. Rhine, the screened open and screened blind matching by Woodruff and George. In the latter the key cards were inverted and unknown; and the cards of the target pack (to be matched against the key cards) were held under the screen by the subject. However improbable it seems, a combination of visual cues from the backs of the key cards and of tactual ones from the cards handled might conceivably be supposed to account for the scoring of the subject.

[2] Rejection of any work from this narrow classification does not imply doubt of the actual exclusion of sensory cues in the work considered. Rather is it highly probable that sensory cues could have had nothing to do with most of these results under the conditions of testing. It is doubtful if any qualified critic would suppose they did. There *might*, however, be a dispute about some of this work because sensory cues are remotely conceivable. The aim back of these narrow criteria is to make sensory cues under the conditions admittedly inconceivable.

Our criteria are most rigorous in ruling out screen tests of BT in which subjects remove the cards *after the call* is made. No sensory cues can be had after the call, and at most, short of actual trickery, with which we are not concerned here, only an accidental touch of the edge or back could be had of the next (uncalled) card. But we can well afford in this survey to be even unnecessarily narrow and strict. The work rejected by this ruling has many valuable features, and like practically all the work referred to as excluded from this survey, it is statistically significant. Carpenter and Phalen conducted two series of screened BT, 8,750 trials, critical ratio 3.9. Also, Woodruff and George's 6,825 trials with a critical ratio of 5.9 is in this discard. In both instances there was tactual contact although it ordinarily came *after* the call.

So far, then, as tactual and visual sensation go, if the cards are simply not in any way or degree seen or touched by the subject during the run, the role of sensory cues is obviously impossible. We have only to read the reports, select all the data meeting the conditions, and the case for and against sensory cues is determined.

3. Only the results of tests made with the ESP cards (or a modification thereof, leaving the probability of success of each trial 1/5) will be included. The ground is again largely that of avoiding debatable issues and keeping the question within easy understanding. That is, the results of tests made with drawings, playing cards, and other objects are, although measurable, not so easily judged; and the reports of the experiments made with other than ESP cards are not so readily available to readers of this JOURNAL.

This provision will appear especially ruthless to those investigators who favor other means of testing than that of using ESP cards. It is admittedly ruthless and is laid down here only to restrict the issue to an area of easy settlement of the single question of sensory cues.

4. Once the above-mentioned criteria are defined, *all* the reported test results that meet them are included in the totals under consideration. Selection *within* the defined class and without the guidance of an acceptable principle is statistically improper and unsafe. But the selection and separate treatment *of* an *a priori* clearly defined class of data *taken as a whole* is not only usual in science but is essential.

The class of data, then, included here and taken in its entirety, regardless of score values, is defined as the score-totals of all of the clairvoyance tests conducted with ESP symbols under conditions in which the subject could have had, according to the reports, no tactual

or visual sensation of the backs, edges, or faces of the cards to be called until after all the calls of the run were made.

* * * * * *

The results to be surveyed fall easily into three general groups, marked by differences in test conditions: Those in which the cards were enclosed individually in sealed, opaque envelopes; those in which portable opaque screens were used to bar vision of the cards; and those in which the cards were in a different room or a different building from the subject and screened from vision by a wall between.

In spite of the narrow criteria defining the evidence acceptable, there are thirteen experimental series eligible, based on the work of 118 subjects investigated by ten experimenters, giving the huge total of 142,825 trials. A number of trials many times as large as this has been rejected as failing in some respect to meet the criteria defined above.

I. CARDS SEALED IN OPAQUE ENVELOPES

The condition described by this heading offers the smallest number of reported tests of the three groupings. Only three researches are involved, but there is a total of 25,225 trials for the three series.

1. For the first series of tests with opaque envelopes we draw from a report published in this number, Dr. Pratt's article on Dr. C. Hilton Rice's work. "Cards sealed in heavy manila envelopes were used, sometimes with an observer and sometimes not. Cardboard shields were placed over the symbols to insure against the possibility that the symbols might be seen through the envelope. There was never any evidence that the sealings had been tampered with." [3]

The following quotation will describe the procedure briefly and give a summary of the results: ". . . Miss T. worked with sealed cards. Sometimes she took the envelopes which Dr. Rice had marked in code and worked alone; at other times she worked as usual in his presence and he checked the results at once from his knowledge of the code. Some of the cards were sealed by himself [Dr. Rice], others in the Parapsychology Laboratory of Duke University. This work extended from September, 1936, until May, 1937. A total of 209 runs were made, giving an average of 5.7." [4] Dr. Pratt's table shows a critical ratio for this series of 5.2.

In similar fashion Dr. Rice tested his subject, A. J., to the extent of 10 runs, obtaining an average of 5.6 per run.

[3] Page 241, this number.

[4] Page 244, this number.

2. In the other and larger opaque-envelope investigation belonging in this classification, that of Price and Pegram with blind subjects, ". . . each card had a blank card glued to the face of it, so that the symbol was inside, and this was sealed in an opaque brown envelope. Each envelope was taped around the outside with gummed paper to make the sealing secure. These cards were identified by a numeral printed on the back of the envelope in a code unknown to the subjects." [5] As to procedure, ". . . . (a) the subjects were never told what symbol card was in any envelope; (b) the code was never revealed to any subject; (c) the code numbers were so small as to be indistinguishable by any subject; (d) three different packs of sealed cards with three different codes were used."[6]

Price and Pegram [7] report among other tests the results of 778 runs [8] with the cards sealed in opaque envelopes given to (partially or totally) blind subjects. They obtained an average of 6.0 hits per run. This gives a critical ratio of 13.1.

In 133 of these runs made by blind subjects the method used was blind matching, BM.[9] In this the key cards were face down and were arranged in an order unknown to the subjects. The target pack was sealed in opaque envelopes. The average reported was 5.95 and the critical ratio was 5.4.

In 329 runs[10] made by the blind subjects with the opaque envelopes the BT method was used with a 9″ x 12″ notebook before the pack as a screen. The average in this series is 6.06, critical ratio of 9.4.

For completeness there must be included 300 trials or 12 runs of 25 which were given my subject, A. J. L., in 1932, using sealed opaque envelopes. The results were exactly what would be most expected by chance—an average of 5.0.[11]

The total of 1,009 runs of both these researches made with cards sealed in opaque envelopes gave an average per run of 5.9 hits. The deviation from the mean chance expectation is 905 and is 14 times the standard deviation.

[5] "Extra-Sensory Perception Among the Blind," JOURNAL OF PARAPSYCHOLOGY, I, 145.

[6] *Ibid.*, p. 148.

[7] *Ibid.*, p. 145.

[8] In order to delete the GESP tests from the pooled totals as published, I have had to consult the authors for some of the figures given here.

[9] See glossary.

[10] Cf. Footnote No. 8, above.

[11] Rhine, J. B., *Extra-Sensory Perception.* Boston: Bruce Humphries, 1934. P. 66.

II. TESTS WITH SCREENED CARDS

Under this subheading are three series totalling 91,700 trials or 3,668 runs of 25. These represent the work of 81 subjects and of six different experimenters. The tests were all made with the cards screened from vision by the subject throughout the run and with no tactual contact with the cards.

1. The first series of this group is likewise taken from Pratt's report of Dr. Rice's work, and consists of 10,475 trials in clairvoyant card calling by two subjects made in part under the observation of Dr. Rice himself, partly by Dr. Pratt, and partly (when Dr. Rice was himself subject) by his secretary, Miss Maude Thweatt. All vision and touching of the cards was eliminated, as the account of the conditions given by Pratt will show: "Like procedure (3), ["In this condition the experimenter alone handled the cards. The subject called out where they were to be placed."] with the additional feature of screening the cards from the subject's sight. The experimenter sat behind and close to a folding screen of opaque cloth 6 feet high. One of the twenty-inch panels was directly between the subject and the cards. The experimenter could see the subject, who sat about 10 feet from the screen, by placing his eye at a narrow crack between the two panels."[12]

The total of 419 runs of 25 average 5.8 hits per 25 and give a critical ratio of 8.0.

2. The second series is made up of 21,500 trials given a single subject by Pratt at Columbia. These trials used a special combination of the BM or blind matching and the STM or screened touch matching procedures. Pratt describes the conditions thus: "As the first step in this direction, the experimenter prepared 28 sets of key-symbol cards in advance. Each set of 5 symbols was shuffled, a blank card was used to cover the face of the bottom symbol to prevent its being accidentally seen at any time, and the 6 cards were fastened together by an elastic band. These 28 sets of key cards were used in rotation, one set being re-shuffled and placed back with the others after being turned face up for checking the results of a series of trials.

"A set of receptacles for the place cards was made by fastening five open shallow boxes to a strip of cardboard. These boxes were just large enough to hold the key-cards, but were deep enough for 50 cards. Five blank cards were provided to cover the face-down key cards after they had been placed in the boxes.

"In arranging the key cards, the following routine procedure was strictly observed. The experimenter first placed the row of boxes

[12] Page 240, this number.

on his knees *under* the table. (Due to the fact that the table was covered with a blanket which hung nearly to the floor, the subject could not have seen the boxes in this position if she had tried.) Then the experimenter selected one of the sets of key cards. Holding them face down under the table, he slid the 5 symbol cards one after another into the 5 boxes. Then the blank cards were placed over the key cards while the boxes were still under the table. Having completed this, the experimenter placed the row of key cards under the edge of the screen [13] before the subject, picked up and cut a deck of shuffled cards, and gave the signal for the subject to begin.

"Her responses consisted in pointing, as previously described, to one after another of the key cards, indicating to the experimenter how to place the cards of the deck to match those before her." [14]

The first 7,800 trials gave a critical ratio of 5.3. Then the subject ceased to score above chance in all her work. She made 13,700 trials more at this level and thus reduced her critical ratio to 3.1; still it is significant.

3. In the third series the 66 subjects were either partially or totally blind. Sixty-one were students in a school for the blind. The investigators, Price and Pegram of the Duke Parapsychology Laboratory, say of their conditions: "(1) No person was allowed to stand beside or behind the experimenter. (2) For all GESP and BT work a stiff-backed notebook (9" x 12") was used as a screen behind and close to which the experimenter held the cards so that the cards were invisible to the subjects and spectators while the calls were being made. (3) The subject was kept at least four feet from the experimenter, usually seated at the opposite side of the table. . . . Two further checks on the results were attempted. (1) Two experimenters worked independently with some of the same selected subjects. (2) A considerable part of the work was witnessed by teachers at the school or was observed and checked by a second experimenter." [15]

Only screened work with the BT method is accepted for this series, and with this method no subject touched the cards either before or after calling. There were all together 39,275 trials made under SBT (screened BT) conditions by these blind subjects, averaging 5.5 per 25 and giving a critical ratio of 10.6.

Of these, 8,225 trials were made with the additional protection

[13] "In this procedure a wooden screen, two feet square, stood upright on the table, separating the subject and experimenter. The lower edge of this screen was raised three inches above the table." JOURNAL OF PARAPSYCHOLOGY, I, 12.
[14] *Ibid.*, 14.
[15] *Op. cit.*, 148.

of coded sealed opaque envelopes enclosing the cards (as described above) with the same screening and other safeguards maintained as with open cards. With this summation of blindness (many subjects totally so), screening, and opaque envelopes, the average was in fact higher, 6.06, and the CR 9.5.

4. MacFarland and George combined the precautions of DT and STM techniques in a series of 20,450 trials by 13 subjects which leave the highest average yet encountered in this review. Their conditions are given in considerable detail:

"In the experimental situation, the subject and the experimenter sat on opposite sides of a table with an upright screen between them. The screen which was used was rectangular in shape, 18 inches high and 24 inches long. It had an opening at the bottom which was 2¾ inches high and 12 inches long. As an additional precaution, a small shield (2½" x 16") was built in front of this opening in such a way that the subject could not possibly see beyond the screen even with eyes brought to the level of the table, and yet both subject and observer were able to see key cards placed in the opening in the screen. Two sets of key cards were used, one normal and the other distorted. Each set consisted of one card of each of the five designs mentioned above. The procedure which was followed consisted of a combination of the STM (screened touch matching) and the DT (down through) techniques. In this procedure the experimenter placed the deck of cards which was to be used face-down on his own side of the table and out of the subject's sight. The key cards were then placed in the opening in the screen. The subject pointed to the key card which he believed matched the first card in the deck, then to the key card which he believed matched the sceond card in the deck, and so on 'down through' the entire deck of 25 cards. The cards in the deck *remained in position* until all 25 responses were completed. This constituted what we refer to hereafter as a 'run.' As the subject pointed to each key card, the experimenter recorded the response on a data sheet prepared for the purpose. The responses were checked and the score recorded at the end of each run. Two decks of normal cards and two of distorted were used throughout the experiment. Normal decks were alternated with distorted in the order of presentation. When the normal decks were used, the key cards were also normal; when the distorted decks were used, the key cards were distorted. Each deck was carefully shuffled after the run in which it was used. At the beginning of the day's work with each subject the decks were again shuffled, except

in those instances referred to later in this report. Although the screen would seem to minimize the importance of the caution, old decks were always exchanged for new ones whenever the old ones became soiled. . . . In all but a very few cases, at least one other observer was present. There was therefore double witnessing of each check-up, as well as of the behavior of the subject throughout the test." [16]

The astonishing critical ratio of 29.2 results from the evaluation of these results. The 20,450 trials averaged 7.1 per run.

The three series just reviewed total 91,700 trials by 81 subjects averaging 5.8 and producing a deviation from mean chance expectation of 3,082, which is 25.0 times the standard deviation.

III. TESTS WITH DISTANCE AND WALLS SEPARATING CARDS AND SUBJECTS

The separation of subject and cards by walls and by distances greater than that which the ordinary across-table ESP test involves is a general condition for the next group of six series of investigations. These were conducted by seven different experimenters on a total of 34 subjects. The six series total 34,225 trials, with an average of 5.7 hits per 25, approximately that of the other two groups of data surveyed. The deviation of the total of the scores from the mean chance expectation of this group of 1,369 runs is + 987, which is 13.1 times the standard deviation.

For lack of a more discriminative principle, that of degree of distance will be used as a basis of order of presentation.

1. The 750 trials reported by Martin, of the University of Colorado, were made with the subject in the room adjoining that in which the experimenter handled the cards. These 30 runs, which averaged 7.8 hits per 25 and give a critical ratio of 7.5, were conducted under the following conditions: "The experimenter shuffled the cards thoroughly, cut them and placed them on the table before her. Exceeding care was taken that the bottom card was visible to no one. The subject then recorded twenty-five guesses on a record blank. The experimenter then read the actual order of the cards to the subject who recorded them. The experimenter carefully watched the recording of each card and was often checked by a third person." [17]

"Series 3. This series of 30 runs or 750 trials was conducted with the experimenter (the author), seated at a table in one room, and the subject seated at a table in an adjoining room. The subject recorded

[16] *Op. cit.*, 94-95. [17] *Ibid.*, 186.

her guesses after receiving a 'start' signal from the experimenter. A door was ajar between the rooms to permit of signals. The order of the pack was recorded by the subject in the presence of the experimenter."[18]

2. A short series of 61 runs or 1,525 trials using DT and BT procedures was conducted by Stuart with H. P. as subject, and a report was included (without much detail because the results did not warrant it) in my semi-popular volume, *Extra-Sensory Perception*. An average of 5.3 was obtained with the non-significant ratio of 1.1.

The distances were from 8 to 30 feet with either one or two walls between cards and subject. The pack was cut by the experimenter before the test and the recording was done by him also, although always in the presence of the subject and commonly with the latter turning the cards over (always close under the experimenter's eye, however).

3. With the help of Miss Ella Phillips Crandall as a second observer, I conducted with the subject E. J. G. 3,725 trials in clairvoyance tests (both DT and BT) at short distances, obtaining an average of 5.7 and a critical ratio of 4.4. I quote from my report as to the conditions: "The calls were recorded by an observer as made, and either at the end of the run of 25 calls or after several packs were called, the check-up would be made. Both subject and observer would check both card and call together, both observing both. No opportunity for error or deception was allowed. During most of the tests, including the best, more than one observer was present. . . . But a still further advance in conditions consists in the separation of subject and cards by distance and a wall. This was done first with a short distance of about 12 to 15 feet, and with one wall, cutting off vision; and second, with 30 feet and two walls. An open door permitted communication, but in some experiments the doors were closed; signalling with a telegraph key was used for the BT condition to indicate when the card had been removed and the call recorded." [19]

4. The report by Warner in this issue comes well within the scope of our criteria. It is hardly necessary to quote at length from his account of conditions since it is easily available. The distance, the double observation and double checking, and other features of the research make it, as its author entitles his report, "A Test Case." His average per 25 trials of 9.3 hits gives a critical ratio of 6.8. This is the highest average of the series here surveyed.

5. The series next in turn has become well known already as the Pearce-Pratt series. It consists of 1,850 trials by H. P. with Pratt

[18] *Ibid.*, 188.

[19] *Character and Personality*, III, 95-96.

as main experimenter and with myself as a second observer in 150 trials. The general average was 7.5. (During the double observation series it was 9.3.) The critical ratio is 10.7.[20] The conditions are given: "The observer and subject synchronized their watches, and arranged to work at a stated time and distance. At the specified time the observer would take the top card from a shuffled pack of ESP cards in the room agreed on and lay it face down on a book in the center of the table without looking at its face. Thirty seconds later the subject in his cubicle in the Duke Library would record a call for the card. At the end of the minute, the observer would remove the card and take the next one. The cards as removed would be kept in order for later recording. Two runs were made per day.

"Groups A, C, and D were made with the observer in the Physics Building and the subject in the Library. The distance was about 100 yards. In group B, the observer was farther away with the cards, approximately 250 yards. The subject was in the same place in the library as in A, C, and D. . . .

"In Groups A-D, the records were sealed up after each sitting and delivered to me before subject and observer got together. In Group D, I was present with Dr. Pratt as a second observer.

"The cards were shuffled between runs, and this shuffling occurred just before using, with H.P. absent from the room. . . . Two packs were used at a sitting."[21]

6. The last and largest series of this group has also the largest distance as a condition. MacFarland, at Tarkio College, has through the medium of the mails supervised 26,125 trials by 30 subjects with distances extending up to 1,400 miles separating subjects from the cards. Reference to the table given in my preliminary report of this work[22] will show that the highest average (5.64) was at a distance of 100 to 300 miles. The general average was 5.5 and the critical ratio 8.3. The following paragraph gives further information: "The tests were made with the DT technique, the cards being kept intact in packs by the experimenter throughout the test period and removed only when checking up. The subjects filled out five columns of a record sheet, one for each pack in the experimenter's desk, at any time they wished on a given day, and sent the sheet to the experimenter to be checked. Double checking was carried out, and the general sponsorship of Dr.

[20] This series was used as the basis of the examination of methods of evaluation made by Greenwood and Stuart (JOURNAL OF PARAPSYCHOLOGY, I, 206). By consulting this as well as the original report further information may be had.
[21] JOURNAL OF PARAPSYCHOLOGY, I, 75. [22] *Ibid.*, 181.

R. W. George, head of the Department of Psychology, was exercised over this series. The cards were well shuffled and kept under careful observation by the experimenter." [23]

The total of the individual trials in Groups I, II, and III, with envelopes, screens, and distance, respectively, is 142,825. The average per run of these 5,713 runs is 5.8. The positive deviation of 4,624 is 30.0 times the standard deviation. The table presents the figures of all series in summary.

TABLE I.

CLAIRVOYANCE TESTS MEETING SPECIAL CRITERIA FOR EXCLUSION OF SENSORY CUES

Experimenters	*No. Subjects*	*Av. per 25*	*Trials*	*Runs*	*Deviation*	*S. D.*††	*C. R.*
I. CARDS SEALED IN OPAQUE ENVELOPES							
P & P	66	6.0	19,450	778	746	56.9	13.1
CHR	2	5.7	5,475	219	159	30.2	5.3
JBR	1	5.0	300	12	0	7.1	0
Total		5.9	25,225	1,009	905	64.8	14.0
Total sealed less screened sealed‡		5.8	16,900	676	555	53.0	10.5
II. TESTS WITH SCREENED CARDS							
CHR & JGP	2	5.8	10,475	419	333	41.8	8.0
MacF & G	13	7.1	20,450	818	1,704	58.3	29.2
JGP	1	5.2	21,500	860†	188	59.8	3.1
P & P	66	5.5	39,275	1,571	857	80.9	10.6
Total	81	5.8	91,700	3,668	3,082	123.5	25.0
III. TESTS WITH DISTANCE AND WALLS SEPARATING CARDS AND SUBJECTS							
LW	1	9.3	250	10†	43	6.3	6.8
DLM	1	7.8	750	30	84	11.2	7.5
CES	1	5.3	1,525	61	17	15.9	1.1
JGP & JBR	1	7.5	1,850	74	187	17.5	10.7
JBR	1	5.7	3,725	149	110	24.9	4.4
MacF & G	30	5.5	26,125	1,045	546	65.9	8.3
Total	34**	5.7	34,225	1,369	987	75.5	13.1
GRAND TOTAL	118**	5.8	142,825*	5,713	4,624	±154.2	30.0

* There are 8,325 trials appearing twice in this table. These are, however, not duplicated in this total. They averaged 6.06 hits per 25.

** These totals do not check with numbers of subjects given because of duplications in some series.

† Equivalent to this number of runs of 25, though actually not so broken up in the experiment.

†† Corrected for Matching Hypothesis; i. e., about 2 per cent larger than for binomial.

‡ Repeated in Group II below.

DISCUSSION

The aim of this survey was to assemble the results of all tests in which there were exceptional assurances that sensory cues were com-

[23] *Ibid.*, 180.

pletely barred by the conditions—barred so effectively that there can be no dispute of the fact. And to this end, many borderline series were rejected for which a case probably convincing to most scientific readers might be made.

The criteria were made so exclusive that the 142,825 trials are not to be suspected of explanation by the hypothesis of sensory cues. Nor are the scores any more readily explainable by chance. The critical ratio of 30.0 represents (if we use the normal probability integral tables as an approximately correct basis of estimate) a probability of only one in 10^{195} that so large a total deviation (as 4,624) would occur by chance alone.

Something has evidently operated in an extra-chance manner in this series of tests reviewed, and under the conditions reported it could not have been a sensory process. But the critical reader must, of course, consider other alternatives as well. Since this treatment is limited to the sensory cue question, it is necessary to be very brief in considering further steps in the weighing of the evidence. Perhaps the most important alternative left for judgment is that of the general reliability and competence of the investigators. For it would appear that if the reports are to be taken with full reliance—even if one of them be so regarded—there is hardly any alternative to the hypothesis that some extra-sensory mode of perception is occurring.

But there is no statistical method for handling reliability of investigators since there is nothing comparable to "mortality tables" for predicting incompetent or unscrupulous experimenters. Nor do we know how to determine how many experimental confirmations of ESP are required for "significance" of investigators' reliability. The next step, then, with these uncertain and non-measurable factors, must therefore be left to the general judgment and inclination of the reader. They are inherent in all research but become prominent when results do not fit in with pre-existent views.

The question of possible dishonesty or of unintentional error on the part of the subjects themselves does not arise when the cards are adequately protected from the senses. The subjects' active participation in the checking was also eliminated in the majority of the series of this survey.[24]

The limited goal of this survey has, I believe, fully justified the discrimination made between data which does and that which does not meet the criteria laid down for the sure exclusion of sensory cues. That

[24] Passive witnessing by the subject is advantageous, but the actual checking and scoring should be wholly in the experimenter's hands.

the criteria are unfairly high for a just decision on some of the rejected data must be clear, but the goal was that of obtaining an undebatable set of criteria rather than a fair one.

With another objective in view, the rating given to various series of test results in this survey would naturally be very different. In fact, some of the best work from the point of view of bearing on the nature of ESP is not quite eligible to this group. A considerable and important series of comparative tests bearing on the character of ESP have been set aside, comparisons (1) of different size and number of stimuli, by L. E. Rhine; (2) of various methods, by Woodruff and George and by Gibson; (3) of screened cards versus unscreened, by Woodruff and George; (4) of intelligence relationships, by Bond; (5) of various drug effects, by myself; (6) of color versus form stimuli, by Carpenter and Phalen; (7) of high-aim versus low-aim calling, by Pegram; (8) of groups versus individual tests, by Sharp and Clark; (9) of telepathy with clairvoyance, in my earlier work; and indeed many other contributions to the growing knowledge of what ESP is in relation to other more familiar processes.

It must be granted, too, that research in telepathy and GESP has not been given fair representation here. Furthermore, it has in general been neglected in late years in actual research for the more simply controlled tests of clairvoyance. But the problems involved are too important to permit this phase of ESP to continue uninvestigated. Telepathy tests can of course be controlled for sensory cues, and I am confident such cues have been excluded in a great part of the research on that branch. The ground for needing to exclude telepathy research in clarifying the issue in this paper does not, fortunately, apply to the research itself.

The standards enforced in this survey are not to be adhered to strictly in judging that work in which the primary aim has been to inquire into the nature of ESP. One may accept with relative confidence, all the more so against the background of the results surveyed here, the exclusion of sensory cues in experiments such as the screened matching tests (with carefully examined cards), screened card calling without the subject touching the cards until after the call, GESP with subjects in separate rooms, and many other conditions in which significantly high total deviations have been found by capable experimenters.

However, it will doubtless follow that in future research in ESP, both in telepathy and in clairvoyance, better safeguarding against "remotely conceivable" sensory cues will be maintained. The means and

methods are more easily available today than ever before. It may appear unusual that all the precautionary conditions now part of the routine were not applied from the start. But in all research the obvious of today was obscure but yesterday.

The primary advance in ESP research, however, has consisted not simply in laying down heavier barriers to sensory cues, but in the development of such ways and means of approaching subjects with a view to measuring the ability in question that they may perform under conditions favorable at the same time both to the ability itself and to the avoidance of experimental weakness. It has been this double objective of the research which has made its progress difficult and slow. But its severest critic must admit that in spite of difficulties it has made some progress.

SUMMARY

In order to remove the question of the adequate exclusion of sensory cues in the ESP research from the range of all dispute, there was made a survey of all results which meet the following very exclusive criteria, designed to admit only data from tests made under conditions in which, as described, sensory cues are undeniably impossible: (1) Clairvoyance tests made with (2) packs of cards with five suits in which the cards were both (3) invisible to and (4) untouched by the subject (5) throughout the run.

There were found 13 series of trials totalling 142,825, made by 118 subjects, with 10 experimenters, and averaging for the 5,713 runs 5.8 hits per 25, giving a critical ratio of 30.0. The odds against so high an average for this length of series resulting by chance alone would be about 10^{195} to one.

It appears logical to conclude that sensory cues were not responsible for the fact that the subjects as a group scored significantly above the mean chance expectation in the ESP tests falling within the scope of this survey.