ESP TESTS WITH ENCLOSED CARDS

J. B. RHINE

ABSTRACT: This report covers 124,275 trials made in ESP tests in which the cards were enclosed in sealed opaque envelopes and boxes. The total score is such as would be expected only about once in 10^{11} times.

About one-third of these tests were checked immediately as usual, and these gave a critical ratio of 13.5. About half of the total trials, however, were made with delay in checking because of geographic distance, and these gave a negative critical ratio of 2.7. Delay in checking appeared inhibitory. This decline effect appeared in six of the seven test procedures used in the research.

The comparison of enclosed card tests (immediately checked) with screened gave results in favor of the enclosed. Enclosing at least offers no physical barriers to ESP, and effectively excludes sensory cues.

A tentative hypothesis is suggested to account for the scoring below chance with the delayed checking and enclosed cards.

INTRODUCTION

IN MEETING the primary requirement of excluding sensory cues in ESP tests, one of the first measures to occur to the experimenter, naturally, is that of enclosing the object to be perceived in a sealed package.

This was the precaution taken in the very first experimentation on the ESP hypothesis at Duke University. But it had long before been done in the first systematic laboratory studies of clairvoyance that were made, those of Professor Charles Richet, French physiologist, in the investigation of his subject, Leonie. In Professor Richet's tests, he enclosed playing cards in supposedly opaque manila envelopes and sealed the envelopes. His results were positive, but unfortunately we cannot on the basis of the report be entirely satisfied with the evidence for the complete opacity of the envelopes. No additional safeguards were added to reduce the possible translucency of the envelopes.

Following Richet's work, sealed envelopes were used in a number of later experiments with persons supposed to possess telepathic or clairvoyant ability. An outstanding case was the test given the Polish engineer, M. Ossowiecki, by Mr. Theodore Besterman, of the Society

for Psychical Research, London,[1] in which a remarkable success was reported.

Likewise in a series of investigations made by Mr. H. F. Saltmarsh[2] of a British medium, Mrs. Elliott, objects enclosed in sealed boxes were used as the basis for the test. In this case, too, the results were reported as successful in the representation of knowledge not available through the senses.

Sealed opaque envelopes enclosing playing cards were used in the tests for clairvoyance conducted by Besterman, Soal, and Jephson,[3] of the Society for Psychical Research, in 1930; but these tests yielded only chance results, in contrast to extra-chance positive scores obtained by Miss Jephson with open cards.[4]

There was, then, no novelty whatever in the introduction of the sealed envelope method of concealing the test object in the early research at Duke, and no reason to regard the procedure as doomed to failure. The results obtained, however, were practically those expected by chance. But during the period when this method of concealing the cards was used, the tests were administered to classes or groups of subjects and were also given (after the manner of Besterman, Soal and Jephson) in very short series of five trials per subject. Whether or not the failure to score above chance was the result of the group testing or of the short series, rather than of the fact that the cards were concealed, was not then determined. Whichever of these conditions, if any, was inhibitory to ESP, at any rate investigation along these lines yielded very little beyond chance results.

At the same time some tentative explorations made with open cards with individual subjects and in longer runs yielded results which were higher in average. While it is now easy to ascribe the difference in results in the two cases to other conditions than that of the enclosing of the cards, it was at that time regarded as possible that there was some inhibitory effect upon the subject due to his having no actual contact with the card he was attempting to call; and so for the time the enclosing of the cards was discontinued.

Somewhat later, when one of the principal subjects was located at some distance from the laboratory, it became necessary to find some method by which we could continue the tests in spite of the distance. The most suitable method seemed to be that of sending him a pack of

[1] *Proc. S. P. R.* 1933, XLI, 345-352.
[2] *Proc. S. P. R.* 1929, XXXIX, 47-184.
[3] *Proc. S. P. R.* 1931, XXXIX, 375-414.
[4] *Proc. S. P. R.* 1928, XXXVIII, 223-71.

cards enclosed and sealed in opaque envelopes, and asking him to call these and return the record of his calls for checking at the laboratory, where a code was kept. The series of tests made with this subject, with the use of these enclosed cards, has been reported.[5] The results gave the number most expected by chance; that is, the average was exactly five hits per 25, in 300 trials.

With this failure of an otherwise successfully scoring subject, the use of the enclosed card method suffered another set-back; and it was some time before it was again introduced. The next attempt, made at the laboratory approximately a year later, with another subject who likewise was scoring successfully on most of the other procedures, also resulted in failure. The score was not appreciably above mean chance expectation in a series of 102 runs through the pack. Soon thereafter another test series of 50 runs made with another outstandingly successful subject led to a considerable deviation on the negative side. Although this was interesting in itself, our experience on the whole with enclosed cards was discouraging in its failure to yield results compatible with the tests made with open cards. If our tests were a fair measure of ESP, then ESP apparently did not work with enclosed cards as we used them.

There were, however, some very good grounds for supposing that the reason for this situation was entirely mental rather than due to the material enclosing the card. A brief summary of the actual safeguarding procedures that had meanwhile come into use in the laboratory will, I think, make this point clear by showing that even more generally obstructive physical barriers had been used already with success.

First, it will be recalled that in the earliest test procedure used with the open cards (BT) the top card of the inverted pack was to be called before it was removed. If the back of the card did not give sensory cues, its inverted position represented an effective concealing of the symbol on the face of the card. The cards used were made of heavy cardboard opaque to a hundred-watt light bulb. For all practical purposes the symbol on the card when lying face down was as good as enclosed in an opaque envelope.

A further step in concealment was taken in the DT procedure, in which the cards were left in position in the pack until the entire 25 calls were made. Thus, with the exception of the top card, only the edges of the cards were visible, and the symbols were thoroughly concealed by an increasingly thick screen as they lay further down in the pack.

[5] Rhine, J. B. Extra-Sensory Perception. Boston: Bruce Humphries, 1934, p. 66.

A still greater step in concealment of the card was made with the use of opaque (usually wooden) screens, behind which the cards were placed while the subject was asked to call them.

The cards were then removed to another room from that occupied by the subject; and finally, in some experiments they were removed to distances involving hundreds of yards and eventually hundreds of miles. With walls and larger distances between the percipient and the object perceived, it is obvious that the conditions offer much more physical obstruction than anything that obtained in any of the enclosed card tests.

It seemed clear, therefore, even at the time, that the limitation in scoring when the cards were enclosed was due to a mental inhibition rather than a physical one. This mental inhibition accordingly presented a problem in itself; and this paper is a result of the investigation of it. The first step was necessarily to ascertain whether or not it was possible for subjects under any circumstances to call, by ESP, cards enclosed in opaque sealed envelopes, and, if so, to relate the results obtained in such an inquiry with the failures mentioned above, and of course if possible with normal open card test results as well.

With a slight anticipation of the substance of the report, it may be added that the first results of this renewal of the use of the enclosed cards led to considerable complication. In some instances subjects who had been scoring well above the chance level with open cards fell definitely below with enclosed, while others did not. Still others dropped, but only dropped to the chance level.

There was, however, considerable variety of conditions used in this exploratory attempt. The variation of conditions and technique did not at first permit of sufficient analysis to enable an understanding to be reached as to the probable factors leading to the complication. In fact, it was not until a considerable portion of the work reported here had been done that a plausible hypothesis for the interpretation of the unusual results was arrived at. It was discovered at length that the tendencies to negative deviation can be grouped under one condition while tendencies to positive deviation are largely confined to the opposite extreme of the same condition. The result is that at least a working hypothesis is offered to explain the earlier mystery of the difficulties with the sealed pack tests. Of its adequacy the reader must judge for himself.

PROCEDURES AND CONDITIONS

The results reported in this paper represent all of the work done with enclosed cards sent out from or used in the Parapsychology Lab-

oratory at Duke University, with the exception of that to be reported by Miss Margaret M. Price later in this Journal.[6] The period covered by these tests extends from 1932 to the present year, although the greater part of the tests has been made since 1935.

The subjects in these tests have in almost all cases been subjects in other test procedures in the ESP research. They include adults and children of both sexes, both blind and seeing, with a wide professional range and a geographical distribution that covers a considerable part of the country. There have been in all 175 subjects, and the experiments have been conducted under the direction of eight investigators. Six of the investigators were psychologists, one a physician, and one an engineer.

Unfortunately for simplicity in the description of the enclosing of the cards, there were several methods used during the period of several years over which this work extended. The changes occurred as a matter of improvement in convenience or added security. However, there is no strain put by the results upon the adequacy of the methods of enclosing. As will be seen later when the enclosed decks were turned over to the subject for use in the absence of an observer, the results were such as would raise no question about sensory cues. On the other hand, when the results were such as might raise the question of alternative hypotheses, in all cases the enclosed cards were handled by the subject in the presence of an experimenter. This would of course have rendered it impossible to tamper with the enclosing without detection.

Two general methods of enclosing were used, depending upon the method used in the tests: For DT purposes whole decks were enclosed in cardboard boxes, and for other test procedures each card was enclosed in a separate envelope.

Most of the boxes were fastened by passing broad strips of gummed paper completely around the box both lengthwise and laterally, overlapping the ends and writing a signature over the end of the strip. There was one instance in which the box was sealed by filling the crevice between cover and bottom with mucilage, thus completely sealing it at every opening between the two parts.

The methods using envelopes passed through a number of stages. All the envelopes, however, were made of heavy manila paper commonly used as coin envelopes. They were just large enough to hold an ESP card with some additional wrapping material. All envelopes were sealed, and both of the sealed flaps were covered by a broad strip of

[6] Miss Price's results are not available at the time of writing and are not different in bearing from the present results.

heavy gummed paper extending over both flaps. In the case of some decks this strip of gummed paper was extended all the way around the envelope.

The first pack of sealed envelopes was made up for easy convenience by wrapping the ESP card in several layers of heavy white paper, then inserting it in the envelope and sealing as stated above. This rendered the envelope opaque to the brightest light but also made a bulky envelope. The next pack was made up by putting a blank card over the face of the ESP card, inserting and sealing both in one manila envelope and inserting this in a second (larger) one which was then sealed as above described. Several packs were made up in this fashion before a third departure led to the gluing of the blank card over the face of the ESP card. The final supply of enclosed cards added to this precaution gummed strips around the pair of cards glued together, the gummed strips running both ways around; that is, lengthwise and laterally. All the above types were opaque to light from hundred-watt bulbs both reflected and direct.

When a pack of cards was mailed to a subject working outside of the laboratory, a point was made of inspecting the pack upon its return for possible tampering. No evidence of any such tampering was discovered. It is recognized, of course, that with the exercise of the fullest available skill a clever person might possibly succeed in opening the envelopes and in sealing them up again so that the opening could not be detected. However, the scores in no instance seemed to warrant the greater labor and precaution required to afford more security on this point.

Care was taken not to leave markings on the face of the envelope which could serve as the basis of associations for the subject. Such associations could under the test conditions at the worst only mislead the subject and were reported by some subjects to be disturbing. If packs became badly worn or marked, they were exchanged.

An identifying number was placed on the back of each envelope, and a code of the symbols was for most packs taken and kept in a code book which was locked in a filing case in my office, the key to which was kept in my possession.

In certain experiments the experimenter was required to memorize the code and to make the check-up immediately following a run. When this was done, either a screen was interposed between the subject and the cards during the checking process, or else the subject was sent out of the room or (in a few cases) required to turn his back completely while the checking was done. The checking was in all respects done

in such a way as to prevent any possibility of the subject's learning the code.

There were three general conditions as to immediacy or delay in checking. There were, first, tests given in which the calls were checked immediately by the experimenter. Second, there were series in which several runs were made before the check-up. And third, there was a large amount of work in which there was considerable delay in the check-up due to the fact that the calls had to be mailed in to the laboratory for checking there.

Under the last two conditions the individual scores for each run were not given the subject. Instead, the average or the total deviation alone was given for a number of runs; and in some series not even the amount of the deviation was mentioned. The results in those cases were simply characterized as "very interesting," or "very good," or in some such language. The direction of deviation was not given. These precautions were taken to render impossible any inference about the code that might be made by any systematic study of the numbered envelopes with such design in mind.

Similar precautions were taken to prevent subjects from identifying cards in the work done for immediate checking. As mentioned above, the subject was not allowed to see the checking process. In addition, the envelopes were kept face upward during the run, which meant that the code number and other identifying marks were not visible. Also, a large portion of the subjects were blind. While many of these were not totally blind, their vision was not capable of perceiving the small symbols on the backs of the envelopes. In spite of their blindness, however, the same precautions were taken.

All the data capable of recheck (that is, of which records of both cards and calls are available) have been rechecked completely or else large enough samples have been taken to insure safety on the question of errors in recording and checking. The extent of these rechecks will be specified following the presentation of results.

Results

Totals. The report covers a total of 4,971 runs of 25, or 124,275 single trials, representing the entire range of tests to date in which ESP cards were enclosed and sealed. A total positive deviation of 947 was produced, and with a standard deviation of ± 141.0 this deviation gives a critical ratio of 6.7. It is evident that with this critical ratio, which represents roughly odds of 10^{11} to one that such a deviation might be

expected by chance, there can be no question of the statistical significance of this work taken in its entirety.

If, then, this evidence is sufficient to warrant the conclusion that something other than chance is operative, and the conditions have been such as to exclude sensory cues, we have in this work further evidence of extra-sensory perception. It is justifiable, then, to analyze further the work represented by these totals.

Scientific method requires that we attempt to associate variations in measurable results with objectively defined conditions under which they were obtained. This will lead us of necessity to deal with sub-series of this total group and to tentative groupings of results with the conditions under which they were obtained. Those apprehensive of unwarranted statistical segregation of data may be comforted, if not by the legitimacy of the analysis itself, at least by the *statistical significance of the total results.*

The following internal comparisons are of interest in these results: First, a comparison of the effect of different periods of delay between calling and checking; second, a comparison of the results obtained with the various techniques of testing, as, for example, matching compared with calling methods; and third, the comparison of scores made with enclosed and with unenclosed cards.

Immediate versus delayed check. The total number of trials falls naturally into three divisions representing different periods of delay in checking. When this division is made it is found that very marked differences in scoring level occur. All those in which the checking was done immediately following the completion of each run are regarded as Division I. Those in which the checking was delayed until the completion of a series (usually five runs; in some instances this involved, however, a delay of some hours) represent Division II. And those in which the results had to be sent in by mail and therefore the checking and reporting to the subject involved a delay of several days, fall naturally into Division III.

Each of these Divisions has a large block of data: Division I, 1,632 runs; Division II, 1,091; Division III, 2,248. *The averages decline as the delay increases.* They are, respectively, 5.69, 5.07, and 4.88. The critical ratios give the very high figure of 13.5 for Division I and the insignificant one of 1.1 for Division II. For the negative deviation of Division III the barely significant figure of 2.7 is obtained. In Table I it will be seen that the individual experimental series run uniformly positive in direction of deviation in Divisions I and II, whereas in

TABLE I

COMPARISON OF IMMEDIATE CHECKING WITH DELAYED

Subject Groups	No. of Subjects	Experimenter	Runs	Dev.	Av.	C. R.
Division I. Immediately Checked						
1. School for Blind	66	Price and Pegram	778	747	5.96	13.20
2. Duke University	17	Rhine and Associates*	251	117	5.47	3.63
3. Montgomery	2	C. H. Rice	223	169	5.75	5.54
4. Columbia	1	J. G. Pratt	222	73	5.32	2.40
5. Grand Rapids	7	E. P. Gibson	118	2	5.02	0.09
6. Oxford Orphanage	13	Pratt and Price	40	24	5.60	1.87
Total of Division I			1,632	1,132	5.69	13.50
Division II. Brief Delay						
1. Duke University	23	Rhine and Associates*	1,039	53	5.05	0.82
2. Grand Rapids	2	E. P. Gibson	52	23	5.41	1.56
Total of Division II			1,091	76	5.07	1.13
Division III. Long Delay						
1. Grand Rapids	20	Rhine and Gibson	619	—103	4.83	2.03
2. New York University	2	Rhine and Sharp	338	+ 31	5.09	0.83
3. Columbia University	1	Rhine and Pratt	335	— 5	4.99	0.13
4. Duke University	19	Rhine and Associates*	379	— 73	4.81	1.87
5. Tarkio College	9	Rhine and MacFarland	365	— 7	4.98	0.18
6. Montgomery	3	Rhine and Rice	212	—104	4.51	3.51
Total of Division III			2,248	—261	4.88	2.69
Grand Total of Divisions I, II, and III			4,971	+947	5.19	6.72

*Chiefly C. E. Stuart, J. L. Woodruff, Margaret H. Pegram, Sara Ownbey Zirkle.

Division III the trend is preponderantly negative for the individual series.

Comparison of test procedures. Most of the test procedures were used in all three of the Divisions just mentioned (I, II, and III), and therefore the comparison of results with the various test procedures can be made in all three Divisions (I, II, and III) and bear out further the effect of delay in checking. It will be seen, for example, that in the main the decline in score level with delay in checking runs through the various subdivisions representing the results of different test procedures. These results are shown in Table II. For example, the scores

TABLE II

COMPARISON OF AVERAGES FOR DIFFERENT TEST PROCEDURES

Division	Checking	No. of Subjects	Number of Runs of 25 and Average Hits for Principal Procedures* Used													
			OM		BM		STM		BT		DT		PDT		DM	
I.....	Immediate..	106	656	5.6	163	5.8	350	5.6	355	6.0	42	5.6	16	5.4	50	5.4
II.....	Brief delay .	25	159	5.2	28	4.8	...	...	123	4.9	120	5.2	12	5.5	647	5.0
III.....	Long delay .	54	888	4.8	23	4.1	344	5.0	248	5.0	467	4.9	32	5.4	246	4.9

*See glossary for brief description of procedures, except PDT and DM. PDT is pre-shuffle card-calling (See *Journal*, II, 1, 47). DM is ESP shuffle or deck matching (II, 2, 120).

for the OM procedure (see glossary) decline for Divisions I, II, and III as follows: 5.6, 5.2, 4.8; for BM, 5.8, 4.8, 4.1. The only outstanding exception to this is in the PDT (pre-shuffle card-calling, test for precognition) in which the numbers of trials are small for comparative purposes. In general, then, it may be observed that the decline with delay in checking is a well distributed effect not confined to any special test procedure.

The comparison of the results of the varying test procedures themselves is a point of primary interest. It may be seen that there is not a high order of difference shown, the greatest difference appearing again with the smaller numbers of trials. But if we omit from present consideration all groups of below fifty runs, the averages within a given Division are fairly close together; here again, too, the three Divisions differ as markedly as before.

While the BT method shows the highest average score, the next highest average is shown by a matching procedure, BM. Due to the fact that the various test procedures used were often chosen by the subject himself in order to provide conditions more satisfactory to him, the various procedures have not been equally represented. This cannot in any case represent then, a thorough and final comparison of these procedures; and no attempt is made to present further comparative statistics. These results do, however, serve the point of general indication of relative success with different procedures; and the point already mentioned, of showing that no single type of error involved (let us say, in recording or checking), would serve to account for all the deviations obtained with a variety of procedures.

Comparison of enclosed and open cards. This comparison is again an incidental one and cannot be made with nice precision. The results obtained with the enclosed cards, however, represent so large a sample that they may profitably be compared to work reported with the same

subjects and experimenters done with ordinary unenclosed cards, even though there was not an exact number of trials done by each subject for each condition. The general averages at least are of interest and value. For purposes of closer comparison we shall use only the results obtained from open card tests in which the cards were entirely screened from the subject or were separated from him by great distance. This will put the two sides of the comparison on an equal footing so far as the exclusion of sensory cues is concerned.

It is necessary to point out, too, that the work with the open and with the enclosed cards was not done in all cases during the same gen-

TABLE III

COMPARISON OF ENCLOSED WITH OPEN CARDS

	Enclosed Card Tests		Open Card Tests	
	Runs	Av.	Runs	Av.
1. School for Blind	778	5.96	911	5.57
2. Montgomery	223	5.75	158	6.04
3. Columbia	222	5.32	312	5.60
4. Grand Rapids	118	5.02	548	5.19
5. Oxford Orphanage	40	5.60	771	5.19
Totals	1,381	5.7	2,700	5.4

eral working period. There is, of course, likely to be a considerable variation on this account. The results are summarized in Table III.

It will be seen that we have no results for comparison with Division III of the enclosed card work, due to the fact that there are no tests comparable for this condition with the use of unenclosed cards. This would have required having unenclosed but screened cards called and the results sent in to the laboratory, with consequent delay.

Division II has no substantial series for comparison in open card tests. A total of 86 runs averaging 5.16, however, was made, alternating with enclosed card tests; but in these runs the cards were not adequately screened.

The comparison of enclosed and open cards, then, is confined to the immediately checked group or Division I, where there are, as mentioned, still many features of inequality in the conditions represented. However, there are 2,700 runs with cards unenclosed-but-screened available for comparison which meet the requirements. They were conducted by the same experimenters with the same groups of subjects. The average for these tests with open cards is 5.4, as compared to 5.7 for 1,381 runs

on the enclosed cards made by the same experimental groups. This difference of 0.3 per run is significant. Greater detail is shown in Table III. One group (No. 5, Duke) of the enclosed card work is omitted because no corresponding open card work was available for comparison.

DISCUSSION

Further evidence of ESP. The primary requirement of statistical significance is met by the results of this investigation. The total 124,275 single trials give a deviation that has a critical ratio of 6.72.[7] The counter-hypothesis of sensory cues is particularly well ruled out by the fact that every card was enclosed in materials unquestionably opaque to any light available to the subject.

A further question should be raised here; namely, whether the conditions excluded the possibility of the subject memorizing the code and identifying the envelopes. The question concerns two different situations: First, that in which the subjects were located geographically at great distances from the laboratory where the code was kept locked up. It would have required collusion with the laboratory staff to obtain the code. One is assured further by the fact that the score averages from all such subjects have been either below the mean chance expectation or very close to it. (See Division III, Table I.)

The other aspect of this question concerns subjects who were tested in the presence of the experimenter. In this instance the only opportunity for the subject to acquire a knowledge identifying the cards in the envelopes lay in his finding visual cues on the front of the envelope and then observing some give-away indication of the experimenter during the checking-up process by which he would identify that particular envelope. Even granting the presence of some identifying visual cues on the envelopes, the other requisite for the identification of the card was definitely lacking, since in the checking-up process the subject was not allowed to look on. In the case of the blind children it can be seen that this alternative too is ruled out by their deficiency as well as by the experimental conditions.

It is next in order to consider the adequacy of the safeguards against errors in handling the data. We find that particularly with the more recent work the evidence is quite reassuring against this hypothesis. First, all of the delayed check work of Division III, except fifty runs, was independently recorded and was capable of independent recheck-

[7] The only enclosed card work done to date not included in this total, that of Miss Price already mentioned, will increase this critical ratio to a still higher figure.

ing. The critical ratio of 2.7 for this block of data is based, then, upon this precaution.

A sample of 433 runs was checked from the original codes. (The average from Division I and II was raised from 5.23 to 5.24; errors of omission are the more common.)

It was possible to recheck 320 runs of the Price and Pegram work with the blind by confirming the arithmetic and making certain that there appeared in the card column for each run exactly five of each symbol. These scores averaged 5.88. Since this average is very close to the total average of 5.96, the sample may be considered an adequate check upon the total results.

Further reassurances on the safeguarding of the checking procedure are as follows: The short series with the Oxford Orphanage group shown in Table I which averages close to the average for Division I was conducted by two experimenters so that double witnessing prevailed throughout. This was true likewise of a unspecified part of the work at the School for the Blind. A further reassurance is offered by the fact that, as Table II shows, the averages for the matching and the calling methods are approximately similar. In view of the very different methods of checking and recording, this would be very unlikely, on a basis of supposition of error, to account for the deviations. Finally, it should be said that in those matching methods for which no independent records of calls and cards were kept, we have eight experimenters responsible for the recording and checking. In the matching procedures the checking is a matter of very simple counting of the number of hits. While error is possible, it is much less technically probable than that of the task of the average physical scientist engaged in reading various balances and meters. Probably in all such instances error occasionally occurs. It cannot, however, reasonably be supposed to have become a habit for so considerable a number of trained professional people as are responsible for these tests. Nevertheless, the reader will have to select according to his own criteria what standard he will accept; it is likely that in some portion of the data he will find that it has been met.

Delay in checking. It is definitely indicated by the averages of Divisions I, II, and III in Table I that a falling off in score average occurred with delay in checking. That this decline is not a mere accident is indicated in a number of ways: First, by the critical ratios of the different subdivisions (these latter were made entirely on the principle of delay in checking); second, there is the fact that, as shown in Table II, even when the test results are further subdivided according

to the test procedure used, there is still shown in all but one subdivision the same decline with delay. It is noteworthy that in the main the subjects who had been scoring high in open card tests and who scored well in the immediate checking dropped lowest in score level in the delayed checking. This can be seen by a comparison of the Montgomery group for Divisions I and III. Also, the lowest scoring subject, H. P., in the Duke University group for Division III was one of the outstanding subjects in tests with open cards. The Grand Rapids group in Division III furnished two outstandingly low scorers. One of these, L. H. G., was the principal high contributor in Gibson's work reported in the December 1937 JOURNAL.

It seems clear that there is some peculiar mental inhibition operating in suppressing the scoring in this delayed check, enclosed card work; that is, if the evidence is sufficient to warrant some speculation. First, it should be noted that mere delay in itself can hardly be regarded as the whole explanation, in view of the fact that, as I reported in this JOURNAL in September of last year, in 101,450 trials conducted with the subjects a long distance from the cards, a significantly positive deviation was obtained in spite of a delay equal to that of the present work with enclosed cards.

But what difference may be supposed to inhere between the distance work with the cards far away from the subject, and the present work in which the cards are enclosed in sealed envelopes and are in the possession of the subject? In both cases the call series has to be copied down and sent to an experimenter by mail and a reply awaited. Both methods presumably are less interesting, or at least less immediately satisfying than the method of immediate checking.

If there is anything about the enclosed card method of delayed checking that will account for the results obtained, it would have to be something which not merely lowered the scoring to chance, but which drove it into a negative deviation; that is, introduced a reversal of the tendency to hit the right card and caused a tendency to be brought into play to miss the cards more often than chance would allow. Presumably this contrary urge or negativism was in all cases entirely unconscious, though on this score the exigencies of courtesy might well have kept a dislike for the enclosed card, delayed check methods from being expressed. One thing is clear, however, that it was not easy to get subjects to take part in any appreciable number of runs with the enclosed card, delayed check method.

But if we suppose a negativistic attitude toward this procedure and condition, what explanation can be furnished for negativism in this case

that would not apply to simple distance tests of clairvoyance with open cards in the hands of the experimenter. At present we have little more than sheer speculation to furnish a working hypothesis. But it appears probable that there is more annoyance given a subject by a sealed package in his possession of which he tries to divine the contents than would be the case if the object were simply far away from him. In the case of the enclosed cards the same envelope is given him time after time, and each time he can only "guess" what it is and is never told; whereas in the open card distance experiment the subject never has to deal with the same card-order over again. It is, I am inclined to suppose, probably the continual frustration of being repeatedly asked to call a pack of cards with no satisfaction ever given as to just how successful a given trial or even a given run has been. It seems thus probable that a positive urge can with sufficient striving, as in the fable of the fox and the grapes, turn sour and become negative. The fact, however, that the subject does not wish to give up the testing which he has promised to carry out may keep him from giving conscious recognition to his annoyance. So much, then, for a merely plausible and wholly untested view of the results of the tests in Division III.

In this connection it should of course be recalled that subjects in ESP tests have been successful in obtaining low scores by deliberate effort at missing the calls. A number of reports of this character have appeared in the literature on ESP, in general the result being that the negative deviation in the low-aim calls approximates that of the positive deviation of the same subject in high-aim calls. The average negative deviation in Division III is not as great as the average positive deviation in Division I, it is true; but even if the speculative hypothesis just proposed to account for the negative deviation in Division III be sound, it need not be supposed that all of the subjects in the tests reported in Division III were negatively oriented. It should be recalled, too, that other negative deviations have been reported without deliberate effort on the part of the subject operating to bring them about. These deviations are only suggestive, but there is a certain obvious relationship to the present work which merits mention.

The first case of negative deviation is that reported by Estabrooks in his series No. 4.[8] This series, in contrast to the other three, which were positive, gave a significantly negative deviation on the calls of the subjects for the suits of playing cards. Thirty-two runs of 20 calls each averaged only 4.06, whereas mean chance expectation is 5. Calls

[8] G. H. Estabrooks, "A Contribution to Experimental Telepathy," *B. S. P. R. Bull.* V, 1937 (Feb.), p. 19.

as to color were also negative. The difference between Series 4 in Estabrooks' report and the other three series, which were positive, is that the subjects were asked, contrary to their expectation, to wait over for this series after they had already finished one. The subjects were students at Harvard College and are represented by Estabrooks as having had no great intrinsic interest in the tests. For example, he was unable to get them back for a retest. One can readily suppose that the process of waiting around for an unscheduled repetition of tests which were not in themselves exciting would stir some at least unconscious impatience and that this might evoke a certain amount of negativism in some subjects.

A similar instance may be found in my report of the early experiments at Duke in the case of the subject, Linzmayer.[9] On this occasion I took advantage of the subject's desire to leave by pre-arranged plan, and persuaded him to stay over for a long series of tests. He did this under sheer pressure of courtesy, as was obvious. His score, which had been averaging approximately 9 hits per 25, dropped to an average of 4 for the next 16 runs.

It is a far jump, but perhaps not without some connection, to pass next to the negative deviations obtained from psychotic patients by Shulman[10] and Van Wiemokly. They found averages of 4.72 and 4.77 for involutional melancholia patients, in which classification there seems to be general agreement that the prevailing attitude of the patient would probably be one of impatience with the tests.

At any rate, it will be of interest to watch for possible cases of negative deviations under conditions in which the subject is to some degree frustrated or led to take a negativistic attitude toward the test. Perhaps some more specific experimental method can be found to evoke this attitude consistently.

Comparison of procedures. The conspicuous thing about this comparison, as has been mentioned already, is the extent of the decline with delay in checking shown by the various subdivisions according to test procedure. It is also noteworthy that the general level of scoring varies very little from one procedure to another within a given Division. In view of the fact that the subjects were as a rule allowed to select their favorite test procedure and yet that there is a substantially large sample of the principal method, it would appear relatively unlikely that any important difference lies between the methods, except as the subject's

[9] Extra-Sensory Perception. Boston: Bruce Humphries, 1934, p. 59.

[10] A Study of Card-Guessing in Psychotic Subjects, J. PARAPSYCHOL., 1938, II, 95-106.

attitude itself becomes determinative. It has been shown by a number of investigators that for given individuals certain procedures appear to be advantageous. A general perspective of these results, however, seems to suggest that these differences are entirely personal and are perhaps matters of taste.

The fact that the total number of runs differs from one subdivision to another might suggest to the critic that a series under a particular heading was discontinued when a favorable critical ratio was obtained. This, however, was not the case, as can readily be understood from the fact that these totals represent summaries of a number of independent investigations in which no investigator knew what the other had done or when the others had discontinued. The total effect, then, may be regarded as quite independent of calculation. In fact, the critical ratios not only for the subdivisions but for the total of the Divisions and of the research itself as a whole were not struck until the occasion of writing this report.

Comparison of enclosed with open cards. Little can be said as a result of this comparison, because of the lack of precise comparative conditions. The primary problem so far as the enclosed card work is concerned was to find out if significant ESP data would be given. Such comparisons as are available are largely incidental; nevertheless they do serve to show, as Table III bears out, that there is certainly no disadvantage in having the cards enclosed in sealed envelopes, provided the subject can be acquainted with his success reasonably promptly and his spontaneous interest in the task thus maintained.

In view of this comparison it is clear that there is no physical obstruction to ESP in the enclosing of the test cards; it is thus all the more certain that if there is, as appears, some reversing tendency induced with the delayed checking of these sealed packs, it must be due to a psychological barrier and not to a physical one.

Summary

A total of 4,971 runs was given a group of 175 subjects under the observation of nine experimenters; and a total positive deviation of 947, with a critical ratio of 6.72, was obtained. The conditions for the exclusion of sensory cues and precautions taken to avoid disclosing the identity of the enclosed cards, along with independent rechecks on the security of the recording and checking, leave no recognized alternative to the hypothesis of extra-sensory perception.

It was found that there was a marked correlation between delay in the checking and decline in the scoring. When the cards were checked

immediately, positive deviations were obtained; and when the cards were sent out to subjects, with the necessity of using the mails for the conduct of the tests, the total score averages became significantly negative. This does not, however, establish that mere delay in checking in itself is sufficient explanation. A tentative hypothesis is suggested to account for the depressed scoring tests involving delayed check.

Comparison of test procedures reveals no outstanding difference in favor of any of the procedures. In six out of seven test procedures used the decline with delay in checking occurred.

An incidental comparison of open card tests with the enclosed card work of the present report gives an average somewhat in favor of the enclosed card work. This incidental comparison is regarded merely as demonstrating that no insuperable physical handicap inheres in the enclosing of the card.

The suggestion is left that where there appears to be a depression of scoring as a result, apparently, of the delayed checking of enclosed cards, this is entirely due to mental rather than physical inhibitions.