

REPLY TO CRITICS OF THE STANFORD EXPERIMENTS ON THOUGHT- TRANSFERENCE

JOHN E. COOVER¹

Late Fellow in Psychical Research, Stanford University

ABSTRACT: Critics of Coover's 1917 monograph, reporting research on thought transference (Stanford University), have stated that conclusions other than those drawn by Coover are indicated by the data. Coover points out that these criticisms are not valid because of fallacious selection of particular experiments, uncontrolled sources of error in the original research, and lack of regard for negative results obtained in experiments in which the subjects were of the "gifted" type.

IN THE MONOGRAPH² that has evoked protests from capable critics there are several sections that relate to the problem of thought-transference: (1) 1000 trials with one subject on the guessing of lotto block numbers, (2) 10,000 trials with normal subjects and 1000 with "sensitives" on the guessing of 40 playing cards, (3) 2400 trials on "the feeling of being stared at," (4) 15,000 trials on the influence of subliminal visual and auditory cues upon guessing, (5) data on mental habits, which play a part in "telepathy" because of community of experience, shown in guessing playing cards and numbers, in reading scientific instruments and in estimating magnitudes, (6) an exhibit of data on inductive probability showing safety of reliance on the theory of probability in evaluating the experiments, (7) a section on infinitesimal probability showing that highly improbable events may be expected to occur in daily life, and (8) a critique of the evidence offered to prove the reality of telepathy. The conclusions were in general negative with respect to thought-transference, but positive with respect to

¹Due to the author's untimely death, this manuscript was left unfinished. I have added only the last part of the analysis of Thouless' criticisms and the concluding statements. The responsibility for these statements rests solely with me, although we were in essential agreement on them.—John L. Kennedy, Fellow in Psychical Research.

²COOVER, JOHN E. "Experiments in Psychical Research at Stanford University" Stanford University Press, 1917, xxiv + 641. (Now out of print.)

the "subjective" experience of telepathy in the laboratory, to the effect of subliminal cues in guessing, to the effect of mental habits and to the trustworthiness of theoretical probability in the evaluation of card guessing experiments.

The criticisms which I shall endeavor to answer apply to the section of experiments with 100 normal subjects on the guessing of 40 playing cards. These criticisms are based upon recomputations of results abstracted from the tables of data exhibited in the monograph, and tend to controvert the negative conclusions reached in the monograph.

Dr. F. C. S. Schiller, in his review³ of the monograph, says of the card-guessing experiments,

The experiments constituted a good test of the hypothesis that the capacity for receiving telepathic impressions is faintly diffused throughout the human race (p. 262).

He examined the final totals of the tables:

	NUMBER RIGHT ON				Guesses
	Card	Color	Number	Suit	
<i>Obtained</i>					
Card-not-Imaged	141	2491	488	1252	4865
Card-Imaged	153	2556	538	1344	5135
<i>Expected by Chance</i>					
Card-not-Imaged	122	2433	487	1216	4865
Card-Imaged	128	2568	514	1284	5135
<i>Deviation</i>					
Card-not-Imaged	19	58	1	36
Card-Imaged	25	11	24	60
Chance Limit	47	150	90	130	5000

and observes,

The results taken in bulk were negative in the aggregate, and so did not support the belief in faint degrees of telepathy. . . . There was a slight excess over the most probable number in both the Card-not-Imaged and Card-Imaged series, and a slight difference in favor of the Card-Imaged experiments in which telepathy might conceivably be operative; but it was not sufficient to warrant any inference as to the reality of telepathy, as it was not above the limits of chance variation.

And he notices that if the totals of the separate thousands of guesses are examined, some of them in the Card-not-Imaged series

³ *Proceedings of the Society for Psychical Research* (London), 1918, 30, 261-273.

might have been taken as better proof of telepathy than any in the Card-Imaged experiments (pp. 262-3).

Thus far the critic is in agreement with the text.

When, however, we come to test the hypothesis that telepathy occurs in certain (*rare*) subjects, the interpretation of these figures is by no means as clear. For it then becomes legitimate to argue that *some* only of the reagents had faint telepathic capacity, and to select their answers for separate analysis. . . . Hence we may select 14 reagents . . . and tabulate their results.

The totals of this tabulation of 14 sets of Card-Imaged experiments and of another tabulation of an equal number of selected results from the Card-not-Imaged series, are as follows:

	GUESSES RIGHT ON		Guesses
	Card	Number	
Card-Imaged.....	54	119	711
Card-not-Imaged.....	49	91	690
Card-Imaged.....	7.6%	16.7%	
Card-not-Imaged.....	7.1%	13.2%	
Chance Limit.....	5.0%	14.8%	

Concerning the Card-Imaged successes, Schiller says,

There was therefore a considerable excess over the ordinary chance distribution. But was it too great to be ascribed to chance, and must it be attributed to an unknown cause ("telepathy")?

He refers to the table of percentages and learns that the per cents right exceed the chance limit, and continues,

This is on the face of it a considerable excess, and if the method of reasoning is sound, would appear to prove some degree of telepathy in the 14 selected subjects (p. 263).

It should be noted that there appears to be some question in Dr. Schiller's mind as to whether this procedure of recomputation is sound.

Dr. Schiller then turns his attention to the Card-not-Imaged results of the 14 subjects who got three or more cards right, and notes that on the "card right" his computation shows them also to exceed the chance limit, and on number right, closely approach that limit. In fact, there is little extra-chance cause to be granted the Card-Imaged group that cannot also be granted the Card-not-Imaged group. Does it seem to be telepathy or clairvoyance that may be the extra-chance cause?

Dr. Schiller notices that in the two groups of 14 best guessers five subjects are common and he computes their results separately:

	GUESSES RIGHT ON		Guesses
	Card	Number	
Card-Imaged.....	19	45	240
Card-not-Imaged.....	22	41	260
Card-Imaged.....	7.9%	18.7%	...
Card-not-Imaged.....	8.5%	15.8%	...
Combined.....	8.2%	17.2%	500
Chance Limit.....	5.5%	15.7%	

The percentages . . . are very markedly higher than the maxima attributable to chance. . . . These figures therefore distinctly point to some source of rightness beyond chance in these cases. As it occurs in the Card-not-Imaged series, it cannot be of the nature of conscious Thought-Transference. But may it not be due to a sort of "lucidity" or clairvoyance in the subject? (p. 264).

Dr. Coover is hardly entitled to deduce from his data that "no trace of an objective thought-transference is found as a capacity enjoyed in perceptible measure by *any* of the individual normal subjects" (p. 124). He should recognize not only, as he does (Coover, p. 65), that his "control" (Card-not-Imaged) experiments assumed the non-existence of any "lucidity," but also that this may have been *wrong*, and that they may not have been "pure guesses," and endeavor to experiment further with the abnormal coincidence of good guessers successful in both series (p. 265).

If the excess of right answers continues (with these subjects), he will be able to establish the existence of telepathy on a statistical basis. Experiments with these five good normal subjects would appear on the evidence to be more promising than with the 10 "sensitive," whose 530 Card-Imaged experiments only showed a comparatively slight excess over probability, and whose guesses in Card-not-Imaged experiments gave no hint of lucidity (p. 265).

(Is it not curious that these selected normal subjects should show more promise for lucidity than do mediums and "sensitives" who are reputed to have it?)

In a footnote (p. 264), Schiller draws attention to the fact that while in the two groups of 14 best guessers, five subjects are common to both sets, in two groups of worst guessers, 24 subjects in the Card-not-Imaged experiments and 20 in the Card-Imaged experiments, only three subjects are common to the two groups. The best guessers tend more than the worst guessers to express their trait in both the Card-Imaged

and Card-not-Imaged experiments. The best guessers showed results above chance expectancy in both groups; the worst guessers in the Card-Imaged experiments showed about chance results in their Card-not-Imaged experiments, but the worst guessers in the Card-not-Imaged experiments showed above-chance expectancy in their Card-Imaged experiments. Thus this last group of worst guessers shares with the best guessers some benefit from imaging the card.

We may carry further this method of recomputation of the results of selected groups:

1. There were 20 worst guessers in the Card-Imaged experiments, who, in sets of approximately 50 guesses, got zero cards right. (This was the basis of their selection.) Total number of trials is 1026.

Card right	= 0	n	= 1026
np expected	= 26	p	= .025
Difference	= 26	q	= .975
D/ σ	= 5.12	σ^2	= 25.0
Chance limit	= 4.24		

Here the difference divided by the standard deviation (σ) of the binomial distribution is 5.12, which exceeds the adopted chance limit of 4.24. Is this proof of reversed telepathy? If the deviation is below the chance limit, the guesser must have known the card in order to miss it.

2. There were 24 worst guessers in the Card-not-Imaged experiments:

Card right	= 0	n	= 1143
np expected	= 28.6	p	= .025
Difference	= 28.6	q	= .975
D/ σ	= 5.41	σ^2	= 27.9
Chance limit	= 4.24		

Again a result which is over the chance limit. Is this reversed clairvoyance?

3. Among a considerable number of sets of drawn cards matched with the *preceding* guess there were seven sets in which three or more cards were right. By selecting these for independent treatment we have the tabulated computation from p (probability of a card right) = .025, n (number of guesses) = 374.

Card right	= 24
np expected	= 9.35
Difference	= 14.65
D/ σ	= 4.85
Chance limit	= 4.24
σ^2	= 9.12

This extra-chance result is significant because it exceeds the limit of chance fluctuation, set high for logical reasons and for convenience at 4.24 times the standard deviation. If the separateness of computation is sound, we seem to have proof of prescience, prediction, or prophecy. The relation is between the guess and the following drawing of a card from a newly shuffled pack.

4. A still more startling result may be shown from the separate treatment of the figures obtained in matching the cards drawn in one set of approximately 50 with those drawn in another; in four such sets there were 4 cards right in each set.

Card right	= 16	n = 197
np expected	= 4.93	p = .025
Difference	= 11.07	q = .975
D/ σ	= 5.03	
Chance limit	= 4.24	
σ^2	= 4.80	

Were this extra-chance result to be warranted, we should have to search for some demon that correlates drawings of cards from immediately shuffled packs manipulated by different persons at different places and times or for mischance in tabulating figures in tables or in marking or counting coincidences; either demon, or (2) experimental, (3) clerical, (4) statistical, or (5) computational error. (I have numbered these alternatives because I believe them to be often neglected when they are potent for accumulating spurious coincidences.)

5. Suppose we now test this method of recomputing results of a group selected from a distribution by selecting for separate analysis only the results of the 8 subjects who got 3 cards right in the Card-Imaged experiments.

n = 399			
Card right	= 24	Number right	= 67
np expected	= 10	np expected	= 40
Difference	= 14	Difference	= 27
D/ σ	= 4.50	D/ σ	= 4.50
Chance limit	= 4.24	Chance limit	= 4.24
σ^2	= 9.73	σ^2	= 35.90

In this last sample, the number of cards right and of number right is also above the limit of chance yet the probability of getting 3 cards right is .093 (the fourth term in the expansion of $(q + p)^{50}$) and $.093 \times 100$ (sets of 50 guesses) = 9.3. On the basis of pure chance, 9 cases of 3 cards right are expected in the distribution. In our distribution of actual results in the Card-Imaged experiments we obtained

8 such cases (one less than chance expectation), but when we select those cases and recompute them separately, we seem to have statistically significant proof of an extra-chance cause.

It is evident that the separate treatment of a part of a distribution calls for a special method. Case 5 shows that part of a chance distribution may be made to show cause apart from chance. All one needs to do is to select sufficient data removed well from the mean of the distribution and he can "prove" the presence of any cause he has an interest in promoting.

Selection for separate treatment should not be made from the figures in the distribution; it should be made from the conditions of the experiment. Or, if selection is made on the basis of figures in the distribution, it should be for the purpose of going back to the data to seek for conditions that may be responsible for the peculiar nature of these figures, rather than for putting them through a statistical mill that is no more applicable to them.

This demonstration of the fallacy of "selected" cases in statistical populations is made for the benefit of those who not only overemphasize but misinterpret Dr. Schiller's criticism. Dr. Schiller, at the beginning of his consideration of the hypothesis that telepathy occurs only in certain rare persons, says,

It then becomes legitimate to argue that some only of the subjects had faint telepathic capacity, and to select their answers for separate analysis (p. 263).

And, after exhibiting that the results of his recomputations were above the chance limits, he says,

This is on the face of it a considerable excess, and *if the method of reasoning is sound, would appear to prove* some degree of telepathy in the 14 selected subjects (p. 263).

He notices that "card right" by the best 14 guessers in the Card-not-Imaged experiments is also above the "chance limit," and then turns his attention to the 5 subjects common to the two groups; after combining their Card-not-Imaged and Card-Imaged scores, he finds them "markedly higher than the maxima attributable to chance," and says these figures therefore *distinctly point to* (not definitely prove) some source of rightness beyond chance in these cases.

May it not be a sort of lucidity in the subject? Doctor Coover is hardly entitled to deduce from his data that "no trace of an objective thought-transference is found as a capacity enjoyed in perceptible measure in *any* of the individual normal subjects" (p. 124). He should recognize not only,

as he does (Coover, p. 65), that his control "Card-not-Imaged" experiments assumed the nonexistence of any lucidity, but also that this may have been *wrong*, and that they may not have been pure guessers.

He does not say that there appear to be excesses here beyond chance guessing and that further experimenting with the best guessers, if they hold up, would "establish the existence of telepathy on a statistical basis" (p. 265), and might be more profitable than further work with sensitives in search of lucidity.

So far, then, (1) Schiller does not demur at the high critical ratio ($D/\sigma = 4.24$) chosen for the limit of chance variation. (2) He agrees that the totals of all the experiments do not pass this limit and that therefore the hypothesis that telepathy as a capacity existing faintly in all men is not supported; he draws attention to excesses beyond chance expectation in the results of selected best guessers that "distinctly point" to some source of rightness beyond chance in these cases. (3) He thinks the experimenter was "hardly entitled" to say that no trace of an objective thought-transference as a capacity enjoyed in perceptible measure by any of the individual normal subjects was found, and (4) he thinks that guessing in the Card-not-Imaged experiments was not pure guessing.

The monograph conclusions were: (1) No telepathy as a common capacity was found because the general totals did not exceed the chosen chance limit; (2) No telepathy as a capacity of single individuals was found because no totals of single sets exceeded the limit of chance. The same conclusions could have been deduced with reference to lucidity. After Schiller's criticism, both of these conclusions still stand. However, something more could have been done with the excesses that challenged Schiller's attention; these are sub-critical, below the limit of chance variation; while they count zero for telepathy, which is required to pass the limit of chance, they may count significantly toward establishing any known causes. But these may wait till their entrance in the next review to be examined.

Dr. Robert H. Thouless, of the Department of Psychology at Glasgow University, in his review⁴ of Professor Rhine's monograph⁵ takes note that Rhine quotes, among other investigations,

that of Coover as providing positive evidence of extra-sensory perception. Since it is a common opinion that Coover's results were entirely negative and show nothing but chance distribution, I have thought it worth while

⁴Dr. Rhine's recent experiments on telepathy and clairvoyance and a reconsideration of J. E. Coover's conclusions on telepathy. *Proc. Soc. Psychical Res.*, 43, 24-37, 1935.

⁵"Extra-Sensory Perception," *Boston Soc. Psychical Res.*, 1934, pp. 169.

to re-examine Coover's figures and will discuss these before proceeding with Rhine's own work. Rhine is undoubtedly right in saying that Coover's results actually show strong evidence against chance. There seem to have been two reasons why Coover himself drew the opposite conclusion: first, he adopted an absurdly high limit for the deviation from mean expectation which might be attributed to chance, and, secondly, he did not consider the possibility that clairvoyance might be active where telepathy was impossible in the Card-not-Imaged series⁶ (p. 25).

Dr. Thouless also inspected the general table presented first above, and stated that

The general tendency of both of these Card-not-Imaged and Card-Imaged series is clearly to exceed mean chance expectation, and in approximately equal amounts. Coover concludes that since the factor of telepathy cannot be present in the first series, the approximate equality of the two groups is due to the fact that the deviations of both are due to chance. A safer conclusion would seem to be that if any factors are present causing deviation from expectation, these are operating in approximately equal amounts in the two conditions of experimentation. At any rate, we shall be justified in lumping the two groups together for statistical consideration. For the remainder of the discussion of these results, I shall do this since it will give us the advantage of the higher significance to be obtained by larger numbers (p. 26).⁷

First, let us consider the reasons for adopting the "absurdly high" limit of chance expectancy. In a footnote on page 85 of Coover's monograph, we find the following explanation:

Although there are certain restrictions to the application of this formula, such as when n (the number of experiments) is small, or when p is very small, its use is peculiarly applicable to the type of data with which we are dealing, and is not unsupported by approved statistical methods already used in the field of psychical research. E., in his "Analysis of Mrs. Ver-rall's Card Experiments" (Proc. Soc. Psychic. Res., 1895, 11, 193-197). Mr. C. P. Sanger used the following formula for the limit of chance deviation from the probable number of occurrences:

$$K = 3\sqrt{2} \cdot (1-q)mq$$

in which q is probability of occurrences and m is number of experiments.

By appropriate transformations of symbols, this formula reduces to the

$$L = \sqrt{2} \cdot 3 \sigma$$

used by Coover to evaluate his card guessing results.

Another hint as to the reasons behind Coover's selection of a high limit of chance variation is contained in a letter written to Mr. Whately Carington in 1938. To quote:

⁶ See Coover (p. 65) for a discussion of assumptions in the Card-not-Imaged series.

⁷ Coover's manuscript ended here.

It is agreed that if there is awareness of objective facts or events, without the functions of the organs of sense, it is of great importance to establish that phenomenon scientifically (i.e., by experiment). Also agreed that no negative judgment is justified now or ever. But what seems not to be agreed is that the history of the experimental work on the problem suggests that there is a very *small* probability that the hypothesis is true, and that therefore a *high* critical ratio for the limit of chance deviation is a logical necessity.

Professor Thouless does not seem to reject this logical assumption as necessary proof of telepathic transmission. He does object to Coover's conclusion that the results were explainable on the basis of chance variation. He says,

The difference between observation and mean chance expectation of the number of cards guessed altogether right by Coover's subjects shows a probability of chance occurrence very much below this limit (50-1. D/σ 2.05), being 200-1 (D/σ 2.58). The existence of some factor favoring correct guessing of the card is strongly indicated. It might be objected that any form of extra-sensory cognition is *a priori* so improbable that we shall be right to insist on a much more severe criterion of significance than we should need, let us say, if we were trying to investigate the difference in fertility of manured and unmanured fields. To this objection there are two replies. First, the question at issue is not, at the moment, whether or not extra-sensory cognition occurred amongst Coover's subjects but whether there was a factor in his experiments favoring correct guessing (such a factor might be some unnoticed error of method). There seem to be no grounds for regarding the presence of such a factor as very improbable. A much more important consideration, however, is that if we are convinced of the *a priori* improbability of extra-sensory cognition, that will be a sound reason for accepting the indications of a 200-1 odds against chance with less conviction than we should otherwise feel; it is no evidence at all for regarding heavy odds against chance as evidence in favor of the operation of chance (p. 27).

The cogency of this criticism is recognized. It remains, however, to return to Coover's description of method in order to evaluate the 200-1 odds. The control of these experiments may be questioned on two points, (1) independence of recording and (2) use of minimal and subliminal auditory cues.

In order to eliminate completely errors in recording at the time of the experiment which are due to knowledge of the correctness of the guess, the recording of the cards and the subject's guesses should be done separately and independently. Coover states:

The experimenter with a watch before him, (1) shuffles the deck of forty playing cards (the face cards being discarded), cuts the pack, and holds

the cards concealed; (2) shakes the dice-box, to determine a control or regular experiment, and, if the latter, the form of content the latter is to have in his mind; (3) if a regular experiment, he turns over the pack, exposing to his view the under card, taps once to signal the reagent that the experimental period begins, holds mental content of card and wills the content to be projected into the mind of the reagent, and, after 15 to 20 or more seconds, taps twice to signal the close of the interval. *After he notes that the reagent has recorded his guess,*⁸ and has turned to his introspections, he records the color, number and suit of the card and the number of the die-spot which conditioned the form of the experiment (as, R5H 1, for Red, Five of Hearts, Die-spot 1—i.e., precisely the same form as the regular, except that the card remained unknown until the reagent had recorded his guess (pp. 53-54).

These conditions for recording would have been completely controlled if the agent could not have seen the subject while the latter was recording his guess.

Referring back to the quotation on method given above, it is apparent that the experiments were carried out in the same room. Although 847 trials were collected with the agent at distances of 1, 2, 3, 4.6, 6, and 10 meters from the percipient, no differential effect to distance was discovered in the analysis of this series. It is possible, however, that occasional involuntary auditory cues may have been operative in producing small excesses over chance expectancy when the results of all the 5,135 Card-Imaged trials were summed. Coover (pp. 452-460) describes methods and exhibits negative results for telepathy or any other factor besides chance variation when conditions of the experiment were such as to exclude sources of error. These experiments were carried out with "sensitives" who claimed to have the ability to obtain information by telepathy.

The following conclusions may be drawn from this discussion:

(1) Critics of Coover's 1917 monograph have confined their attention to the card guessing series with normal subjects, attempting to show that selected subjects from this group of 100 gave evidence of ability to produce extra-chance scores. In his discussion of Schiller's criticisms, Coover has demonstrated the statistical fallacy in such selection.

(2) The critics have not given due regard to the negative results in his tests, using 10 professional and private "psychics," of the hypothesis that only a few selected cases have telepathic ability.

⁸ Italics mine. J. L. K.

(3) Uncontrolled sources of error may explain the positive deviations on whole card right in the "normal series" of card guessing results. Coover's experimental conditions did not completely eliminate the possibility of errors in recording and involuntary auditory cues.

(4) These considerations do not support statements concerning Coover's "positive" findings for telepathy and clairvoyance which are appearing in the modern literature of "Extra-Sensory Perception."