

Paranormal Communication *“Error Some Place!”*

by Charles Honorton

*Review of the ESP controversy
traces debate from statistical
and methodological issues to
the a priori critique and the
paradigm of “normal science.”*

Asked his opinion of ESP, a skeptical psychologist once retorted, “Error Some Place!” I believe he was right, but for the wrong reasons. Western science has always been ambivalent toward the mental side of reality, and it is perhaps not surprising that the occurrence of “psychic” phenomena is one of the most controversial topics in the history of science.

The first serious effort toward scientific examination of psi claims was undertaken by the Society for Psychical Research (SPR), founded in London in 1882 for the purpose of “making an organized and systematic attempt to investigate the large group of phenomena designated by such terms as mesmeric, psychical, and spiritualistic.” The SPR leadership included many distinguished scholars of the period, and similar organizations quickly spread to other countries, including the American Society for Psychical Research, founded in New York in 1885 under the aegis of William James, who himself took an active role in early investigations of mediumistic communications.

These turn-of-the-century investigators focused much of their attention on authenticating individual cases of spontaneous experiences suggestive of psi communication. While a great deal of provocative material was carefully examined and reported (e.g., 13), the limitations inherent in the case study approach prohibited definitive conclusions. However thoroughly authenticated, spontaneous cases cannot provide adequate assessment of such potential sources of contamination as chance coincidence, unconscious inference and sensory leakage, retroactive falsification, or deliberate fraud.

Charles Honorton is director of research in the Division of Parapsychology and Psychophysics, Department of Psychiatry, Maimonides Medical Center, Brooklyn, N.Y.

Early experimental approaches primarily involved the "telepathic" reproduction of drawings at a distance (62). While often striking correspondences were obtained, the experimental conditions did not usually provide for random selection of target (stimulus) material, and were not always totally adequate with respect to the possibility of sensory leakage, intentional or otherwise.

Neither the spontaneous case studies nor the early experimental efforts made much impact upon the scientific community, though they drew critical comment from prominent period scientists. "Neither the testimony of all the Fellows of the Royal Society, nor even the evidence of my own senses," proclaimed Helmholtz, "would lead me to believe in the transmission of thought from one person to another independently of the recognized channels of sense." Thomas Huxley declined an invitation to participate in some of the early SPR investigations, saying he would sooner listen to the idle gossip of old women.

*The rudiments of an experimental methodology
for testing psi were suggested three
centuries ago by Francis Bacon.*

In *Sylva Sylvarum*, a work published posthumously, Bacon discussed "experiments in consort, monitory, touching transmission of spirits and forces of imagination." He suggested that "the motions of shuffling cards, or casting of dice" could be used to test the "binding of thoughts. . . . The experiment of binding of thoughts should be diversified and tried to the full; and you are to note whether it hit for the most part though not always" (2).

The application of probability theory to the assessment of deviations from theoretically expected chance outcomes was introduced to psychical research in 1884 by the French Nobel laureate, Charles Richet, in experiments involving card-guessing. The popularity of card-guessing as an experimental methodology was greatly influenced by the work of J. B. Rhine and his associates at Duke University in the early 1930s. Rhine (50) devised a standard set of procedures around a simplified card deck containing randomized sequences of five geometric forms (circle, cross, wavy lines, square, and circle). These "ESP cards" were prepared in packs of 25, and each "run" through the pack was associated with a constant binomial probability of 1/5, since subjects were not given trial-by-trial feedback. Providing the experimental conditions were adequate to eliminate illicit sensory cues, recording errors, and rational inference, statistically significant departures from binomial chance expectation were interpreted as indicating extrasensory communication.

Initially, "telepathy" tests consisted of having a subject in one room attempt to identify the order of the cards as they were observed by an "agent" in another room. In "clairvoyance" tests, the subject attempted to "guess" the order of the cards directly, as they lay concealed in an opaque

container or in another room, without an agent. "Precognition" tests, introduced somewhat later (59), required the subject to make anticipatory guesses of the card order before the pack was shuffled or otherwise randomized.

Rhine introduced the term "ESP" in his first major report on the Duke University work in 1934 (50). He reported a total of 85,724 card-guessing trials, carried out with a wide variety of subjects and under a wide range of test conditions. The results as a whole were astronomically significant, though informal exploratory trials were indiscriminately pooled with those carried out under more carefully controlled conditions. The best-controlled work during this period was the Pearce-Pratt distance series of clairvoyance tests (58), in which the subject, Pearce, located in one building, attempted to identify the order of the cards as they were handled, but not viewed, by Pratt, the experimenter, located in another building. The level of accuracy obtained in this series of 1,850 trials was associated with a probability of 10^{-22} .

As a stimulant to experimental research, Rhine's work had unprecedented influence. For the first time a common methodology was adopted and employed on a large scale by a number of independent and widely separated investigators. For the first time, also, the scientific community was confronted with a body of data, collected through conventional methods, which it could no longer ignore—nor too hastily accept. The wide-scale adoption of the card-guessing methodology was accompanied by a plethora of critical articles, challenging almost every aspect of the evaluative techniques and the experimental conditions. During the period between 1934 and 1940, approximately 60 critical articles by 40 authors appeared, primarily in the psychological literature. While card-guessing is no longer the primary methodology in experimental parapsychology, the questions which arose over its use are of equal relevance to the more sophisticated approaches used today.

The first major issue concerned the validity of the assumption that the probability of success in the card-guessing experiments was actually 1/5.

If chance expectation is other than $1/5$, the significance of the observed deviations would obviously be in doubt. This issue was quickly resolved by mathematical proof and through empirical "cross-checks," a form of control series in which responses (guesses) were deliberately compared with target orders for which they were not intended (e.g., responses on run n_1 matched with the target sequence for run n_2). Empirical cross-checks were reported for 24 separate experimental series involving a total of 12,228 runs (305,700 individual trials). While the actual experimental run scores (e.g., guesses on run n_1 compared to targets for run n_1) were highly significant and yielded a mean scoring rate of $7.23/25$, the control cross-check

scores were in all cases nonsignificant, with a mean scoring rate of 5.04 (43).

Several critics questioned the applicability of the binomial distribution as a basis for assessing the statistical significance of ESP card-guessing data. Willoughby (78) proposed the use of an empirical control series, but later withdrew the suggestion after comparing the two methods (79). Alternative methods of deriving the probable error and recommendations for using the empirical standard deviation were also proposed and later withdrawn (21, 22). Concern over this issue diminished and was generally abandoned following the publication of a large chance control series involving half a million trials and demonstrating close approximation to the binomial model (12).

Another question arose about whether the binomial model provides sufficient approximation to the normal distribution to allow use of normal probability integral tables for determination of significance levels (17). Stuart and Greenwood (73) showed that when the normal distribution is used as an approximation to the binomial model, discrepancies are important only with cases of borderline significance and few trials.

The use of the binomial critical ratio (z) to evaluate the significance of the ESP card-guessing deviations was generally approved by professional statisticians (6, 20). Fisher (10), however, commented that high levels of statistical significance should not be accepted as substitutes for independent replication. In another vein, Huntington (20) asked, "If mathematics has successfully disposed of the hypothesis of chance, what has psychology to say about the hypothesis of ESP?"

The most frequently expressed methodological concern was the possibility of some form of "sensory leakage," giving the ESP subject enough information about the targets to account for significant, extrachance results.

As early as 1895, two Danish psychologists, Hansen and Lehmann (16), reported that with the aid of parabolic reflectors subjects could detect digits and other material silently concentrated upon by an agent. In these experiments, the subject and agent sat with their heads close to the foci of two concave mirrors. While the agent concentrated on the number, he made a special effort to keep his lips closed. Under these conditions, the subjects were frequently successful in identifying the number. These results were interpreted by Hansen and Lehmann as supporting the hypothesis of "involuntary whispering." The utilization of subtle sensory cues was demonstrated in a careful investigation by S. G. Soal of a stage "telepathist" (66). There were also reports, such as the case of "Ilga K.," a mentally retarded Latvian child who could read any text, even in a foreign language, when someone stood behind her, reading "silently." Experiments with dictaphone recordings revealed that "Ilga" was responding to very slight auditory cues (3).

It is clear that at least some of the early exploratory series reported in Rhine's monograph were open to criticism for inadequate controls against sensory cues. While Rhine did not base major conclusions on such poorly controlled data, inclusion of them in his monograph provided a ready target for critical reviewers and sidetracked discussion away from the better controlled work, such as the Pearce-Pratt series, which was not susceptible to explanation by sensory cues.

Defects in an early commercial printing of ESP cards were reported by several investigators (18, 25). It was found that the cards were warped and could under certain conditions be identified from the back. This discovery circulated widely for a time as an explanation of all successful (i.e., statistically significant) experimental series. The parapsychologists retorted that defective cards had not been employed in any of the experiments reported in the literature and that, in any case, they could not account for results from studies involving adequate screening with such devices as opaque envelopes, screens, distance, or work involving the precognition paradigm in which the target sequences were not generated until after the subject had made his responses (53, 54, 72).

By 1940 nearly one million experimental trials had been reported under conditions which precluded sensory leakage. These included five studies in which the target cards were enclosed in opaque sealed envelopes (41, 45, 46, 54, 59), 16 studies employing opaque screens (7, 8, 11, 19, 33, 34, 35, 38, 41, 42, 44, 45, 46, 59, 71), ten studies involving separation of subjects and targets in different buildings (50, 51, 52, 53, 34, 32, 8, 77, 61, 60), and two studies involving precognition tasks (59, 75). These data are summarized in Table 1. The results were independently significant in 27 of the 33 experiments. By the end of the 1930s there was general agreement that the better-controlled ESP experiments could not be accounted for on the basis of sensory leakage.

The hypothesis that significant "extrachance" deviations in ESP experiments might be attributable to motivated scoring errors was investigated in several studies. In one investigation (26), 28 observers recorded 11,125 mock ESP trials. Of these, 126 (1.13 percent) were misrecorded. Observers favor-

Table 1: ESP card-guessing experiments (1934-1939) excluding sensory cues^a

Method	Studies	N (Trials)	Mean/25	p <
"Clairvoyance" paradigm, stimuli in sealed, opaque envelopes	5	129,775	5.21	4.0×10^{-11}
"Clairvoyance" paradigm, stimuli concealed by opaque screens	16	497,450	5.44	2.0×10^{-12}
Distance ^b	10	164,475	5.37	10^{-12}
Precognition paradigm ^c	2	115,330	5.15	2.95×10^{-6}

^a References given in text.

^b Includes work with both "telepathy" and "clairvoyance" paradigms

^c Stimuli generated after subjects made their responses

able to the ESP hypothesis made 71.5 percent more errors of commission (increasing ESP scores), while those who were unfavorable to the ESP hypothesis made 100 percent more errors of omission (decreasing ESP scores). Murphy (37) reported an analysis of 175,000 trials from experiments reporting positive evidence for ESP and found only 175 errors (0.10 percent). Greenwood (12) reported only 90 recording errors in rechecking his 500,000-trial control study, of which 76 were errors of omission.

Some critics also alleged that improper selection of data could account for experimental successes. This could be done in several ways: (a) selection of subjects; (b) selection of particular blocks of data out of larger samples; (c) selection of one of several forms of analysis; and (d) selective reporting of particular studies. The questions raised have sometimes been stated cynically in the form, "Parapsychologists must run 100 subjects before they find one with 'ESP'." As if in defense against this charge, a number of the reported studies specifically stated that all of the data collected were included in the analysis (see 43, pp. 118-124, Table 12).

Concerning selection of subjects, Warner (76) suggested two criteria: first, results of "poor" subjects must be included up to the point when they are discontinued since it does not matter how many trials a given *subject* makes as long as all of the *trials* (for all subjects) are included; second, exclude *all* preliminary trials (for both "good" and "poor" subjects) and use preliminary screening studies to select "good" candidates for formal work. These criteria were generally endorsed by the chief critics of the period (e.g., 23).

The question of *post hoc* selection of analyses was not a point of serious concern in the period between 1934 and 1940, though it is relevant to the assessment of some of the process-oriented investigations reported more recently. The question of whether nonsignificant studies were withheld from publication involves an issue which is of great concern to the behavioral sciences as a whole (70, 81) and one which is difficult to accurately assess since there is no way of knowing how many studies may have been withheld from publication because their results failed to disconfirm the null hypothesis.

Several studies of American Psychological Association publication policies (4, 70, 81) indicate that experimental studies in general are more likely to be published if the null hypothesis is rejected at the conventional .05 and .01 alpha levels than if it is not rejected. These studies also indicate that a negligible proportion of published studies are replications. Bozarth and Roberts (4), in a survey of 1,334 articles from psychological journals, found that 94 percent of the articles involving statistical tests of significance reported rejection of specific null hypotheses; only eight articles (less than 1 percent) involved replications of previously published studies.

With respect to the implications of such selection for the ESP hypothesis, there are two partial answers. First, considering the degree of critical interest which prevailed in the 1930s, it seems unlikely that nonsignificant findings would have been repressed during this period; second, the high levels

of significance attached to some of the reported ESP investigations would necessitate postulating astronomical numbers of "chance" trials in order to dilute the overall deviations to chance. To take one example, consider the Pearce-Pratt series of 1,850 trials which yielded $p = 10^{-22}$. As Soal and Bateman (66) pointed out, it is difficult to believe that 10^{10} (ten thousand million) sets of 1,850 trials could have possibly been carried out between 1934 and 1940 (or, for that matter, since 1940). But, as Soal and Bateman suggest, ". . . if we posit this absurd estimate as an upper limit [with overall chance totals], that would still give us odds of 10^{10} . . . against the supposition that the Pearce-Pratt results were a run of pure luck."

The possibility of obtaining significant "extrachance" results by stopping an experimental series at "favorable" points was also raised (9, 31). While this "optional stopping" hypothesis was generally agreed to be of significance only in cases of marginally significant results, it led to the adoption of several procedural modifications: specification of the total number of trials in advance of data collection, or accumulation of data in blocks of predetermined size.

The possibility was raised by several critics that hand-shuffled cards may display a tendency to stick together or otherwise produce patterns which could produce spurious results (24, 82). While the cross-check type of control series, described earlier, failed to reveal any evidence of patterning, there was a general trend away from hand shuffling in the later published studies, which utilized tables of prepared random numbers as a basis for generating target sequences.

There was—and is (e.g., 15)—a rather widespread belief that most of the evidence supporting the ESP hypothesis originated in the Duke University studies and that most independent replications by other investigators were nonconfirmatory. A survey of the published literature between 1934 and 1940 fails to support this claim. Table 2 shows all the published experimental reports during this period which provided statistical treatment of the data. Inspection of this table reveals that a majority (61 percent) of the outside replications report significant results ($p < .01$) and that the proportion of significant studies was not significantly greater for the Duke University group ($\chi^2 = 1.70, 1 df$).

By 1940, the active methodological controversy was over.

The issues raised were, for the most part, legitimate, and investigators modified their procedures to safeguard their results from methodological criticism. The major issues raised since 1940 center on alleged anomalies in probability theory and the hypothesis of widespread investigator fraud.

Spencer Brown (68, 69) has suggested that statistically significant card-guessing studies provide evidence, not of extrasensory modes of communication, but of fundamental defects in probability theory. He makes three criticisms of random number sequences: (a) published random number

Table 2: Breakdown of experimental ESP studies (1934-1939)

	N (Studies) ^a	N studies reported significant ($p < .01$)	% signif.
Duke group	17	15	88
Non-Duke	33	20	61
Total	50	35	70

^a Includes all English-language studies involving assessment of statistical significance of data, 1934-1939 inclusive.

χ^2 (Duke vs. non-Duke \times significant vs. nonsignificant) = 1.70 (1 df)

sequences have been “doctored” prior to publication in order to remove certain nonrandom features; this practice, according to Spencer Brown, makes such sequences nonrandom and invalidates the use of standard significance tests; (b) the source of some random number sequences involves randomizing machines which utilize the unpredictability of human behavior when examined for microscopic variation; such variation, says Spencer Brown, may be predictable enough to account for observed anomalies in random sequences, as well as some of the significant results reported in ESP guessing experiments; (c) Spencer Brown produces evidence to show that anomalous (significant) departures from probability theory can be obtained by matching columns of random numbers (39).

A detailed examination of these points was undertaken by Scott (64). With respect to “doctored” sequences, Scott showed that the maximum error due to rejected (edited) sequences would not affect interpretations of results which are more than marginally significant and could, in fact, increase the likelihood of making a Type II error. On the hypothesis that ESP results are due to some kind of hyper-regularity affecting both the target sequence and the response (guess) sequence simultaneously and similarly, Scott makes the point that this would lead to the expectation of similar results from matching *any* set of humanly produced random sequences. The cross-check type of control series and the Greenwood chance control series described earlier demonstrate that this is not the case. The anomalies reported by Spencer Brown (68), obtained by arbitrarily matching columns of random numbers, have been criticized on the basis of *post hoc* selection (40) and illustrate not that there are fundamental defects in probability theory, but rather that significant deviations from chance can occur in any data where hypotheses and analyses are not specified in advance.

The most recent phase of the ESP controversy centers on the hypothesis of investigator fraud. This argument was most forcefully presented in a lead article in *Science*, entitled “Science and the Supernatural,” by G. R. Price (47), who began with the following observations:

Believers in psychic phenomena . . . appear to have won a decisive victory and virtually silenced opposition. . . . This victory is the result of an impressive amount of careful experimentation and intelligent argumentation. . . . Against all this evidence, almost the only defense

remaining to the skeptical scientist is ignorance, ignorance concerning the work itself and concerning its implications. The typical scientist contents himself with retaining . . . some criticism that at most applies to a small fraction of the published studies. But these findings (which challenge our very concepts of space and time) are—if valid—of enormous importance . . . so they ought not to be ignored.

Following Hume's argument on miracles, Price asserted that ESP is "incompatible with current scientific theory," and that it is therefore more parsimonious to believe that parapsychologists cheat than that ESP is a real phenomenon. He concluded, "My opinion concerning the findings of the parapsychologists is that many of them are dependent on clerical and statistical errors and unintentional use of sensory clues, *and that all extra-chance results not so explicable are dependent on deliberate fraud or mildly abnormal mental conditions*" (47, p. 360). This extraordinary critique and the ensuing discussion in *Science* (5, 36, 48, 55, 56, 65) were widely reviewed. As Meehl and Scriven (46) pointed out, Price's argument rests on two highly questionable assumptions, namely that contemporary scientific knowledge is complete, and that ESP necessarily conflicts with it. Seventeen years later, in an "Apology to Rhine and Soal," Price retracted his accusations of investigator fraud (49).

Very similar arguments have, however, been made more recently by the British parapsychological critic C. E. M. Hansel (14, 15), who began his examination of the ESP hypothesis by suggesting that "the *a priori* arguments . . . may even save time and effort in scrutinizing the [ESP] experiments. . . . In view of the *a priori* arguments against it we *know in advance* that telepathy, etc., cannot occur."

Because of the "*a priori* unlikelihood" of ESP, Hansel's examination of the literature centered primarily on the possibility of fraud, by subjects or investigators. He reviewed in depth four experiments which he regarded as providing the best evidence of ESP: the already-mentioned Pearce-Pratt distance series (59); the Pratt-Woodruff (44) series, also conducted at Duke; and Soal's work with Mrs. Stewart and Basil Shackleton (66), as well as a more recent series by Soal and Bowden (67). Hansel showed, in each case, how fraud *could* have been committed (by the experimenters in the Pratt-Woodruff and Soal-Bateman series, and by the subjects in the Pearce-Pratt and Soal-Bowden experiments). He gave no direct evidence that fraud was committed in these experiments but said, "If the result could have arisen through a trick, the experiment must be considered unsatisfactory proof of ESP, *whether or not it is finally decided that such a trick was in fact used*" (15, p. 18, italics mine).

Hansel's argument is unclear, inasmuch as he quite properly insists that no single experiment can be conclusive, then proceeds to show that none is, given the theoretical possibility of fraud by subjects or investigators. Hansel's only conclusion after more than 250 pages of careful scrutiny was that these experiments were not "fraud-proof" and therefore not conclusive proof of ESP.

Two recent examples, one involving cancer research (74) and the other involving parapsychology (57), serve to remind us of the importance of cross-validation in the assessment of *any* experimental finding. In both cases, it should be added, the fraudulent acts were detected in-house, by the researchers themselves. The point is that in the final analysis an experimental finding is of value and is to be taken seriously only to the extent that it leads to further inquiry. To regard any experiment as an end in itself is to remove it from the domain of experimental science. It is obvious that hypothetical construct, such as ESP, cannot be validated by any isolated experiment, no matter how well controlled it might be. Independent replication is a necessary prerequisite.

***The claim that psi phenomena operate
outside the framework of physical probability
has been a major source of a
priori arguments against acceptance of ESP.***

It has been suggested that to accept ESP requires the rejection of physics. This is absurd, and it is worth noting that such arguments have usually been advanced and defended by psychologists rather than by physicists.

The debate over the incompatibility of physics and ESP has been conducted almost exclusively within the framework of nineteenth-century deterministic physics, wherein the ultimate constituent of physical reality was still believed to be solid matter. Inasmuch as modern microphysics has exorcised the material out of matter and deals with processes which on our macrophysical level of sensory perception are every bit as erratic and anomalous as ESP, the *a priori* claim that ESP violates specifiable laws of physics can no longer be considered to be of more than historical interest.

ESP and other psi phenomena, while no longer incompatible with physics, are not yet accounted for by physics; but then, neither are the more familiar processes of memory and conscious experience. Indeed, the transformation of "raw feels" into conscious experience is no less a problem for the neurophysiologists of today than it was for the speculative philosophers of classical antiquity. Sir John Eccles, among others, has repeatedly warned, "We should not pretend that consciousness is not a mystery."

The ESP controversy illustrates several features of the paradigmatic view of science developed by Thomas Kuhn (28). Normal science, according to Kuhn, is essentially a clean-up operation, constrained by a broad theoretical framework, or paradigm, which defines the boundaries of legitimate inquiry. Paradigms are scientific world views which provide coherence and structure and determine the types of questions to be posed of nature, as well as the manner in which answers are sought. Normal science is thus a process of paradigm-articulation, rather than of discovery. Within the paradigm structure of normal science, observations which conflict with

the paradigm are seldom made; anomalies are ignored. When the anomalies become sufficiently persistent that they can no longer be ignored, they are hotly disputed. Eventually, a new paradigm is tentatively erected which attracts a group of adherents, and a period of crisis ensues which Kuhn calls a paradigm clash.

In this review I have focused at some length on the period of the 1930s, not because it provides the best available evidence for ESP or the best understanding of the processes underlying its operation—it does neither, but rather because it was during this period that the major *substantive methodological issues* were raised and to a large extent consensually resolved. Since 1940, well over 10,000 journal pages devoted to parapsychological research have been published, and at least 250 experimental studies have been reported. The methodological foundations of the research have gradually diversified, enlarging and enriching the scope of inquiry and providing a basis for more sophisticated study. Automated testing equipment has replaced card-guessing in forced-choice ESP tasks, and quantitative methods have been developed for the objective assessment of psi interactions in nonguessing tasks. Psychophysiological techniques, permitting determination of psi-optimal organismic states, have been introduced and utilized in conjunction with experimental methods more closely approximating the conditions under which psi interactions occur *in vivo*. More important, parapsychological investigators have to a large extent shifted their attention away from the "proof-oriented" approach, which can only reaffirm the presence of anomaly, toward systematic attempts to identify the antecedent conditions necessary for the occurrence and detection of psi interactions, the delineation of positive attributes, and the study of individual differences. Only through the pursuit of such "process-oriented" research can we ever hope to achieve the goals of control, assured replicability (or at least predictability), and eventual understanding.

REFERENCES

1. American Psychological Association. "ESP Symposium at the American Psychological Association." *Journal of Parapsychology* 2, 1938, pp. 247-272.
2. Bell, M. "Francis Bacon, Pioneer in Parapsychology." *International Journal of Parapsychology* 6, 1964, pp. 199-208.
3. Bender, H. "The Case of Ilga K.: Report of a Phenomenon of Unusual Perception." *Journal of Parapsychology* 2, 1938, pp. 5-22.
4. Bozarth, J. D., and R. R. Roberts. "Signifying Significant Significance." *American Psychologist* 27, 1972, pp. 774-775.
5. Bridgman, P. W. "Probability, Logic, and ESP." *Science* 123, 1956, pp. 15-17.
6. Camp, B. H. "Statement under 'Notes.'" *Journal of Parapsychology* 1, 1937, p. 305.
7. Carpenter, C. R., and H. R. Phalen. "An Experiment in Card Guessing." *Journal of Parapsychology* 1, 1937, pp. 31-43.
8. Crumbaugh, J. C. An Experimental Study of Extra-sensory Perception. M.A. Thesis, 1938, Southern Methodist University Library.
9. Feller, W. "Statistical Aspects of ESP." *Journal of Parapsychology* 4, 1940, pp. 271-298.
10. Fisher, R. A. Letter to J. B. Rhine, cited in (50), p. 45.

11. Gibson, E. P. "A Study of Comparative Performance in Several ESP Procedures." *Journal of Parapsychology* 1, 1937, pp. 264-275.
12. Greenwood, J. A. Analysis of a Large Chance Control Series of ESP Data." *Journal of Parapsychology* 2, 1938, pp. 138-146.
13. Gurney, E., F. Myers, and F. Podmore. *Phantasms of the Living* (2 vols.). London: Trubner and Co., 1886.
14. Hansel, C. E. M. "Experiments on Telepathy in Children." *British Journal of Statistical Psychology* 13, 1960, pp. 175-178.
15. Hansel, C. E. M. *ESP—A Scientific Evaluation*. New York: Scribners, 1966.
16. Hansen, F. C. C., and A. Lehmann. "Über Unwillkürliches Flüstern." *Philosophical Studies* 17, 1895, pp. 471-530.
17. Heinlein, C. P., and J. H. Heinlein. "Critique of the Premises and Statistical Methodology of Parapsychology." *Journal of Psychology* 5, 1938, pp. 135-148.
18. Herbert, C. V. C. "Experiment in Extra-sensory Perception." *Journal of the Society for Psychical Research* 30, 1938, pp. 215-218.
19. Humphrey, B. M., and J. A. Clark. "A Comparison of Clairvoyant and Chance Matching." *Journal of Parapsychology* 2, 1938, pp. 31-37.
20. Huntington, E. V. "Is It Chance or ESP?" *American Scholar* 7, 1938, pp. 201-210.
21. Kellogg, C. E. "Dr. J. B. Rhine and Extra-sensory Perception." *Journal of Abnormal and Social Psychology* 31, 1936, pp. 190-193.
22. Kellogg, C. E. "New Evidence (?) for Extra-sensory Perception." *Scientific Monthly*, 1937, pp. 331-341.
23. Kellogg, C. E. Critical comments in (43), pp. 228-238.
24. Kennedy, J. L. "A Methodological Review of Extra-sensory Perception." *Psychological Bulletin* 36, 1939, pp. 59-103.
25. Kennedy, J. L. "The Visual Cues from the Backs of the ESP Cards." *Journal of Psychology* 5, 1938, pp. 149-153.
26. Kennedy, J. L., and H. F. Uphoff. "Experiments on the Nature of Extra-sensory Perception: III. The Recording Error Criticism of Extra-chance Scores." *Journal of Parapsychology* 3, 1939, pp. 226-245.
27. Kennedy, J. L. Critical comments in (43), pp. 218-220.
28. Kuhn, T. S. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press, 1962.
29. Lemmon, V. W. "The Role of Selection in ESP Data." *Journal of Parapsychology* 3, 1939, pp. 104-106.
30. Lemmon, V. W. Critical comments in (43), pp. 222-228.
31. Leuba, C. "An Experiment to Test the Role of Chance in ESP Research." *Journal of Parapsychology* 2, 1938, pp. 217-221.
32. MacFarland, J. D. "Discrimination Shown Between Experimenters by Subjects." *Journal of Parapsychology* 2, 1938, pp. 160-170.
33. MacFarland, J. D., and R. W. George. "Extra-sensory Perception of Normal and Distorted Symbols." *Journal of Parapsychology* 1, 1937, pp. 93-101.
34. Martin, D. R. "Chance and Extra-chance Results in Card-matching." *Journal of Parapsychology* 1, 1937, pp. 185-190.
35. Martin, D. R., and F. P. Stribic. "Studies in Extra-sensory Perception: I. An Analysis of 25,000 Trials; and II. An Analysis of a Second Series of 25,000 Trials." *Journal of Parapsychology* 2, 1938, pp. 23-30; 287-295.
36. Meehl, P. E., and M. Scriven. "Compatibility of Science and ESP." *Science* 123, 1956, pp. 14-15.
37. Murphy, G. "On Limits of Recording Errors." In (1), pp. 262-266.
38. Murphy, G., and E. Taves. "Covariance Methods in the Comparison of Extra-sensory Tasks." *Journal of Parapsychology* 3, 1939, pp. 38-78.
39. Oram, A. T. "An Experiment with Random Numbers." *Journal of the Society for Psychical Research* 37, 1954, pp. 369-377.
40. Oram, A. T. Correspondence. *Journal of the Society for Psychical Research* 38, 1955, pp. 143-144.

41. Pratt, J. G. "Clairvoyant Blind Matching." *Journal of Parapsychology* 1, 1937, pp. 10-17; and "The Work of Dr. C. Hilton Rice in Extra-sensory Perception, pp. 239-259.
42. Pratt, J. G., and M. Price. "The Experimenter-subject Relationship in Tests for ESP." *Journal of Parapsychology* 2, 1938, pp. 84-94.
43. Pratt, J. G., J. B. Rhine, B. M. Smith, C. E. Stuart, and J. A. Greenwood. *Extrasensory Perception After Sixty Years*. Boston: Bruce Humphries, 1966. (Originally published in 1940 by Henry Holt, New York.)
44. Pratt, J. G., and J. L. Woodruff. "Size of Stimulus Symbols in Extra-sensory Perception." *Journal of Parapsychology* 3, 1939, pp. 121-158.
45. Price, M. "A Comparison of Blind and Seeing Subjects in ESP Tests." *Journal of Parapsychology* 2, 1938, pp. 273-286.
46. Price, M., and M. Pegram. "Extrasensory Perception Among the Blind." *Journal of Parapsychology* 1, 1937, pp. 143-155.
47. Price, G. R. "Science and the Supernatural." *Science* 122, 1955, pp. 359-367.
48. Price, G. R. "Where is the Definitive Experiment?" *Science* 123, 1956, pp. 17-18.
49. Price, G. R. "Apology to Rhine and Soal." *Science* 175, 1972, pp. 359.
50. Rhine, J. B. *Extra-sensory Perception*. Boston: Bruce Humphries, 1964. (Originally published in 1934.)
51. Rhine, J. B. "Telepathy and Clairvoyance in the Normal and Trance States of a Medium." *Character and Personality* 3, 1934, pp. 91-111.
52. Rhine, J. B. "Some Selected Experiments in Extra-sensory Perception." *Journal of Abnormal and Social Psychology* 31, 1936, pp. 216-228.
53. Rhine, J. B. "The Question of Sensory Cues and the Evidence." *Journal of Parapsychology* 1, 1937, pp. 276-291.
54. Rhine, J. B. "ESP Tests with Enclosed Cards." *Journal of Parapsychology* 2, 1938, pp. 199-216.
55. Rhine, J. B. "Comments on 'Science and the Supernatural.'" *Science* 123, 1956, pp. 11-14.
56. Rhine, J. B. "The Experiment Should Fit the Hypothesis." *Science* 123, 1956, p. 19.
57. Rhine, J. B. Comments. *Journal of Parapsychology*, 1974, p. 38.
58. Rhine, J. B., and J. G. Pratt. "A Review of the Pearce-Pratt Distance Series of ESP Tests." *Journal of Parapsychology* 18, 1954, pp. 165-177.
59. Rhine, J. B., B. M. Smith, and J. L. Woodruff. "Experiments Bearing on the Precognition Hypothesis: II. The Role of ESP in the Shuffling of Cards." *Journal of Parapsychology* 2, 1938, pp. 119-131.
60. Riess, B. F. "A Case of High Scores in Card Guessing at a Distance." *Journal of Parapsychology* 1, 1937, pp. 260-263.
61. Shulman, R. "A Study of Card-guessing in Psychotic Subjects." *Journal of Parapsychology* 2, 1938, pp. 96-107.
62. Sinclair, U. *Mental Radio, with a Preface by Albert Einstein*. Springfield, Ill.: Charles C. Thomas, 1962. (Originally published in 1930.)
63. Shapiro, A. "Review of ESP—A Scientific Evaluation by C. E. M. Hansel." *International Journal of Clinical and Experimental Hypnosis* 16, 1968, pp. 133-134.
64. Scott, C. "G. Spencer Brown and Probability: A Critique." *Journal of the Society for Psychical Research* 39, 1958, pp. 217-234.
65. Soal, S. G. "On 'Science and the Supernatural.'" *Science* 123, 1956, pp. 9-11.
66. Soal, S. G., and F. Bateman. *Modern Experiments in Telepathy*. New Haven: Yale University Press, 1954.
67. Soal, S. G., and H. T. Bowden. *The Mind Readers: Recent Experiments in Telepathy*. New Haven, Conn.: Yale University Press, 1954.
68. Spencer Brown, G. "Statistical Significance in Psychical Research." *Nature* 172, 1953, pp. 154-156.
69. Spencer Brown, G. *Probability and Scientific Inference*. New York: Longmans, Green, 1957.
70. Sterling, T. C. "Publication Decisions and Their Possible Effects of Inference Drawn

- from Tests of Significance—Or Vice Versa." *Journal of the American Statistical Association* 54, 1959, pp. 30-34.
71. Stuart, C. E. "The Effect of Rate of Movement in Card Matching Tests of Extra-sensory Perception." *Journal of Parapsychology* 2, 1938, pp. 171-183.
72. Stuart, C. E. "A Review of Certain Proposed Hypotheses Alternative to Extra-sensory Perception." *Journal of Abnormal and Social Psychology* 33, 1938, pp. 57-70.
73. Stuart, C. E., and J. A. Greenwood. "A Review of Criticisms of the Mathematical Evaluation of ESP Data." *Journal of Parapsychology* 1, 1937, pp. 295-304.
74. *Time*, April 29, 1974.
75. Tyrrell, G. N. M. "Some Experiments in Undifferentiated Extra-sensory Perception." *Journal of the Society for Psychical Research* 29, 1935, pp. 52-68.
76. Warner, L. "The Role of Luck in ESP Data." *Journal of Parapsychology* 1, 1937, pp. 84-92.
77. Warner, L. "A Test Case." *Journal of Parapsychology* 1, 1937, pp. 234-238.
78. Willoughby, R. "Critical Comment: The Use of the Probable Error in Evaluating Clairvoyance." *Character and Personality* 4, 1935, pp. 79-80.
79. Willoughby, R. "A Critique of Rhine's 'Extra-sensory Perception.'" *Journal of Abnormal and Social Psychology* 30, 1935, pp. 199-207.
80. Willoughby, R. Critical comments in (43), pp. 220-221.
81. Wolins, L. "Needed: Publication of Negative Results." *American Psychologist* 14, 1959, p. 598.
82. Wolfe, D. L. "A Review of the Work on Extra-sensory Perception." *American Journal of Psychiatry*, 1938, pp. 943-955.

