Journal of Parapsychology, Vol. 56, June 1992

CRITIQUE OF THE PEAR REMOTE-VIEWING EXPERIMENTS

BY GEORGE P. HANSEN, JESSICA UTTS, AND BETTY MARKWICK

ABSTRACT: The Princeton Engineering Anomalies Research (PEAR) program has produced a number of experimental reports discussing remote-viewing results and analyses. This work is reviewed with attention to methodological and statistical issues. The research departs from criteria usually expected in formal scientific experimentation. Problems occur with regard to randomization, statistical baselines, application of statistical models, agent coding of descriptor lists, feedback to percipients, sensory cues, and precautions against cheating. Many of the issues of remote-viewing methodology were identified by Stokes and Kennedy over 10 years ago. It is concluded that the quoted significance values are meaningless because of defects in the experimental and statistical procedures.

In the last 10 years, the Princeton Engineering Anomalies Research (PEAR) laboratory has conducted extensive research on remote viewing. This work has been reported in a number of publications, including Dunne, Jahn, and Nelson (1983); Jahn and Dunne (1987); Jahn, Dunne, and Nelson (1987); Nelson, Jahn, and Dunne (1986). Briefer accounts can be found in Jahn (1982); Jahn and Dunne (1986); R. G. Jahn, Dunne, and E. G. Jahn (1980, 1981); Jahn et al. (1983); Nelson and Dunne (1987). The most recent report (Dunne, Dobyns, & Intner, 1989) includes results of 336 formal trials and constitutes the single largest database of remote viewing that has been reported in some detail.

Surprisingly, this research received little attention in five lengthy critical works on parapsychology (Alcock, 1990; Druckman & Swets, 1988; Hansel, 1989; Hyman, 1989; Palmer, 1985). It is also noteworthy that the work has been given virtually no coverage in two parapsychology textbooks (Edge, Morris, Rush, & Palmer, 1986; Nash, 1986). Palmer (1985, p. 57) defended his omission saying: "As procedural details of the subsequent trials are not included in the report, a methodological critique cannot be undertaken." Although the report (Dunne et al., 1983) contained 178 pages, some might consider Palmer's statement to be reasonable. Only pages 3–7 (dou-

An earlier version of this paper was presented at the 34th annual conference of the Parapsychological Association in Heidelberg, Germany, in August, 1991.

ble-spaced) are devoted to procedure; most of the report consists of tables and graphs.

Since the release of Palmer's report, the PEAR remote-viewing experiments have been presented in three refereed journal articles (Jahn & Dunne, 1986; Jahn, Dunne, & Nelson, 1987; Nelson, Jahn, & Dunne, 1986). This research is becoming frequently cited; for instance, Braude (1987) described it as "careful work" (p. 573). McConnell (1987, pp. 204–205) approvingly quoted the results. T. Rockwell and W. T. Rockwell (1988) said: "They have avoided questions of sensory leakage in remote perception" (p. 361), "and they have carried the statistical analysis rigorously to the ends of any possible argument we can envision" (p. 362). They made similar statements in a popular magazine, (T. Rockwell & W. T. Rockwell, 1989). Nelson and Radin (1987) specifically promoted the work as meeting Alcock's (completely erroneous) statistical objections. Beloff (1989, p. 365) declared their statistical procedure to be "a valid method of assessment." Popular authors Alexander, Groller, and Morris (1990) and Ellison (1988) have also promoted the research. Scott (1988) briefly critiqued the work, and a more detailed comment on the methodology, statistics, and reporting now seems appropriate.

A Brief Overview of the PEAR Remote-Viewing Trials

An up-to-date summary of the PEAR remote-viewing work is given by Dunne, Dobyns, and Intner (1989). They report results for a total of 411 trials, 336 of which are designated as "formal." We will summarize the protocol for the 336 formal trials as presented in their report.

A percipient (receiver) attempts to describe an unknown geographical location where an agent (sender) is, has been, or will be situated at a specified time.[1] In each case the agent and the percipient were known to each other. The date and time of the geographical target site visitation were specified in advance. In the "volitional" mode (211 trials) the agent was free to choose the target, whereas in the "instructed" mode (125 trials) the target site was randomly selected without replacement from a pool of potential targets. Different series generally used different target pools. During a trial, the agent spent about 15 minutes immersed in the scene, consciously aware of the intent of the experiment. The percipient usually selected a convenient time, sometimes several days before or after the specified target visitation, and unmonitored, recorded perceptions by writing, drawing, or, occasionally, by tape-recording them.

For most of the trials (the 277 ab initio trials) the first step of the analysis was to have the agent and the percipient each give yes/no responses to a "descriptor" list of 30 questions. The responses to these questions were intended to describe the target location. The original 59 formal trials were conducted before the list had been developed and thus were coded by independent judges (ex post facto).

In order to assess the quality of a match between a target and a response in this procedure, a score is calculated for each trial. PEAR developed 5 different scoring procedures. Most recently they have focused on their Method B, which we discuss here. For the computation, weighting factors (alphas) were used, where each alpha was the proportion of target sites for which the descriptor question was answered affirmatively. The numerator of the score was created by adding 1/alpha(i) if question i was answered correctly as "yes" by the percipient, and 1/[1 − alpha (i)] if question i was answered correctly as "no" by the percipient. The denominator was calculated by adding these terms as if all 30 questions had been answered correctly (the highest possible score).

The next step in the analysis was to create an array of "mismatch" scores by matching descriptor lists for targets and responses from different trials. There were 327 unique targets in all of the series. The 106,602 scores for the "universal empirical chance distribution" were derived by matching each of those targets with the first response generated for each of the other 326 targets (327 × 326 = 106,602).

The mean and standard deviation of this "universal empirical chance distribution" were .5025 and .1216, respectively. The shape of the empirical distribution approximated the normal curve. Therefore, z scores for each correctly matched target–response pair were derived by computing the score for the match and then standardizing it by subtracting .5025 and dividing the result by .1216. These calculations were repeated for several subsets of trials in order to do comparisons (e.g., instructed vs. volitional target selection). For such comparisons, the score for a trial was computed using the alphas derived from the subset in which the trial occurred.

[1] PEAR titles their procedure "*precognitive* remote perception" (emphasis added); however, a substantial number of trials involved retrocognitive or real-time remote viewing.

The overall results for the formal trials gave a composite z of 6.355, with a corresponding p value of 1.04×10^{-10}

Methodological Problems

Randomization

Randomness is the foundation upon which all statistical inference is built. Without any source of randomness, it is impossible to assign probabilities to events. In ESP experiments, subjects' responses cannot be considered random in any sense, and it has therefore long been recognized that targets must be randomly selected. In fact, telepathy research was one of the very first areas in any science to adopt formal randomization (Hacking, 1988). Stanford reports that random selection is "regarded as a sine qua non of adequate ESP-test methodology" (Dr. Stanford replies..., 1986, p. 14). Morris (1978, p. 51) states: "Targets must be selected randomly, rather than by any human decision or fixed set of rules, because such rules and decisions are inevitably patterned or biased and thus are either potentially inferrable by the subject or else may inadvertently match a similar pattern or bias in the subject's responses." This is not merely a hypothetical issue. One of us (Hansen) has observed a number of ganzfeld sessions with naïve subjects who did not understand that targets were randomly selected.[2] During the judging phase, several subjects specifically commented that they were considering picking the judging pool item that would most appeal to the sender (believing the sender had been the one to select the target).

A subject's preferences and biases are likely to shift over time and may vary with factors like moods. For instance, someone who is depressed might be less likely to notice bright colors and activity. Since the agent is completely free to choose the target in the volitional mode, the percipient might be able to infer characteristics of the selected target from knowledge of the agent's mood. Because the targets are geographical locations, other factors might be inferred by the percipient. For example, if the weather was cold or rainy, the agent might choose an indoor location in order to remain comfortable. Similarly, an agent who was aware of the likes and dislikes of the percipient could choose a target to maximize the chances of a match.

[2] The subjects were indeed earlier told that the target was randomly selected, but there is a great amount of information for new subjects to absorb. They usually don't understand all aspects of the procedure the first time they take part.

Dunne et al. (1989) report that 211 of 336 formal trials (63%) were in the "volitional" mode. For these, there was no random selection of the target whatsoever. This is virtually unique in modern parapsychology. For the remaining 125 trials, only in the most recent report is there any information as to how the random selection was made. The information provided is very limited and does not meet the reporting guidelines recommended by Hyman and Honorton (1986).

Agent Coding of the Descriptor Lists

As explained above, PEAR uses a descriptor checklist of 30 items to define a geographical site for statistical evaluation. Nelson et al. (1986) report that "encoding of the target is normally performed by the agent at the time of visitation" (p. 279). This coding is a subjective procedure and can introduce artifacts even though the target *site* is randomly selected. For instance, Descriptor 6 reads: "Is any significant part of the scene hectic, chaotic, congested, or cluttered?" This is vague, and open to interpretation. Houck (1986, p. 36), using a similar descriptor list, noted: "People simply do not answer the questions about the same scene in the same way." An agent sensitive to a percipient's predispositions or current mood might very well code vague descriptors to be consistent with the coding anticipated to be done by the percipient on the given day.

This issue can be conceptualized as a problem of nonrandom target selection. For purposes of statistical evaluation, the actual target is the encoded descriptor *list*. The encoding of the list is done by the agent, who consciously selects descriptors (even though the target *site* may have been randomly determined). This flaw has long been understood, and there are several variants of it. Materials produced by the agent at the time of the trial (or later) should not be used in the judging process. Stokes (1978a, 1978b) noted that if photographs of the target site were taken at the time of sending, it would be inappropriate to use them in the judging procedure. Photographs could reveal which aspects of the site the agent found most interesting. Kennedy (1979a) also raised the issue. Both commentators made the point directly in regard to Dunne's early remote-viewing work, which is now included in the PEAR database (in the "ex post facto" category). Schlitz and Gruber (1981) published a rejudging and reanalysis of one of their own experiments because information from the agent was provided to the original judges; the recomputed p value was less extreme than the first result (by a factor of about 300). Humphrey, May, and Utts (1988, p. 382) noted that "in an actual experimental series, it is critical that the target fuzzy sets be defined *before* the series begins. Because of the potential of information leakage due to bias on the part of the analyst, it is an obvious mistake to attempt to define the target fuzzy set on a target-by-target basis in real time." Their footnote added: "This is actually a quite general comment that applies to *any* descriptor list technology" (emphasis in the original). Problems of this sort are very easily avoided. The simple, straightforward remedy, as noted by Humphrey et al., is to encode all potential target sites before the experiment begins.

Shielding of Agent from Percipient

Although Dunne et al. (1989, pp. 4–5) state that "precautions are taken to ensure that perceptions are recorded and filed before percipients have any sensory access to target information," the report contains no details as to how the agent and percipient were kept apart (in order to convincingly exclude sensory leakage between the two). This is of concern because it is reported that "percipients usually select the time and place most convenient for them to generate their descriptions, and no experimenters or monitors are present during the perception period" (Dunne et al., 1989, p. 4). Virtually no precautions are described that would preclude the agent's and percipient's coming in contact, or a third party's coming into contact with each of them, and innocently conveying information about the target. Needless to say, this is not usual procedure in parapsychology.

Potential Cheating by Subjects

Deception by subjects has long been a problem in psi research (Hansen, 1990), and experimenters have an obligation to guard against trickery. Because of the lack of methodological controls described above, there is a potential for cheating. When the target is selected by the agent, there may be collusion between the agent and percipient. Further, the protocol allows cheating by one partner acting alone. One may try to consciously match the response biases of the other. Morris (1982) specifically addressed the consequence of nonrandom target selection in relation to cheating. He wrote: "Unless selected randomly from an equally attractive target pool, targets are likely to have certain sensible, preferable characteristics that

would allow a psychic familiar with whomever chose the target to infer rationally the nature of the target" (p. 21). PEAR's methods made it easy to cheat. Without the use of randomly selected targets and adequate shielding of the agent from the percipient, it is virtually impossible to detect even simple trickery. Parapsychologists have long been aware of such problems and have issued strong warnings. Rhine and Pratt (1957/1962) wrote: "With GESP and pure telepathy, precautions have to be elaborate and have to be adapted to the special needs of the experimental situation. This methodological problem is often taken too lightly; as we have said, GESP is the hardest psi-test procedure to control adequately against error, especially error due to deception" (p. 37).

We should point out that we have no reason to think cheating actually took place in the PEAR research. Dunne et al. (1989) noted that positive effects come from a large portion of the subjects. However, according to their Tables E and F, Subject 10 contributed 77 trials as percipient and 167 as agent, for a total of 244 trials (i.e., over 70% of the formal trials). Because the procedures allow deception by either percipient or agent acting alone, the contribution of that subject should be considered. If we remove Subject 10's trials from the set, the z score drops from 6.355 ($p = 1.04 \times 10^{-10}$) to 2.17 ($p = .015$, one-tailed).[3]

Statistical Issues

Free-response ESP statistical issues are quite complex and involve many subtleties. The reader unfamiliar with the topic may wish to consult Utts (in press) for an overview as well as the other references we cite. In this section we will give rather brief descriptions of the problems as they relate to the PEAR analyses. Here we restrict our discussion to use of PEAR's preferred Method B for calculation of scores. In the Appendix we describe an optimal-guessing strategy that could artifactually inflate scoring for Method A. As we demonstrate, the potential inflation is severe.

Stacking

The database contains instances of stacking, that is, a single target with multiple responses (see Thouless & Brier, 1970, for a dis-

[3] For purposes of the computation we assumed, as did PEAR, that the trials were independent. We recognize this to be an invalid assumption.

cussion). Such responses cannot be considered statistically independent. This problem has been recognized from the early card-guessing experiments (e.g., Goodfellow, 1938), and Child (1978) has given a good illustration of how it can affect a free-response situation. Random selection of *targets* is necessary (but not sufficient) for statistical independence. Dunne et al. (1989) report that 120 of the 336 formal trials involved multiple percipients. They do give a breakdown of results for trials with single and multiple percipients. However, for the overall result, the 120 trials are considered to be individual and independent; there appears to be no correction for the nonindependence.

Dependence Due to Target Selection Method

For the "Instructed" trials, in which the target site is randomly selected, it is stated that "the target is determined by random selection, *without replacement*, from a pool of potential targets prepared in advance" (emphasis added) (Dunne et al., 1989, p. 13). The "without replacement" procedure determines that the trials are not statistically independent, and this must be accounted for in the statistical method. The PEAR analysis failed to take it into consideration. In the Appendix we show that this and related flaws can result in a p value incorrect by several orders of magnitude.

A further problem develops if the percipient is given trial-by-trial feedback in a "without replacement" sampling regime (Stokes, 1981). As Kennedy (1979b) pointed out, the percipient might systematically avoid describing features of previous targets. This can introduce a severe artifact (for a discussion see Hyman, 1977, 1988). This is a serious problem in the PEAR analysis because the mismatch distribution is used to assess the significance of the correct matches. If a percipient avoids descriptors in one trial based on feedback from previous trials, the mismatch scores could be artifactually deflated.

Target Pool Definition

Target pools with varying characteristics were used in different series. In the case of a randomly selected target from a prespecified pool, the subjects should not be able to infer the identity of the target, but might have some idea about the range of targets in that particular pool. For instance, they might know that the target possibilities lie in the Princeton area and not, say, on the plains of

North Dakota. Responses to the pools in the two areas may be quite different and may reflect the available information. The analyses conducted by PEAR often lump targets from a variety of pools together for computing mismatches. Thus, in the analyses, responses are compared with some targets that were not available for actual random selection for that trial. This can introduce an artifact in the baseline.

PEAR's Position

In the summer of 1988, the PEAR group was given a manuscript written by Hansen which raised many of the points discussed in this paper. Perhaps the Dunne et al. (1989) technical report was an attempt to address those points as well as criticisms raised by Scott (1988). A primary strategy of PEAR's report was to compare subgroups. Sometimes this was done against the global baseline distribution, and sometimes new baseline distributions were created by mismatching target and response pairs within the subgroups. For instance, they did evaluate the results of separate agent-percipient pairs. However, such comparisons do not answer the numerous problems outlined above. The difficulties are inherent in the methodologies and statistical models they used.

The Dunne et al. (1989) technical report also contains a number of invalid statistical arguments. For example, their Figure 1 and the accompanying discussion are designed to show that, if matched scores are compared to local mismatch distributions, then encoding biases cannot artifactually inflate the z scores. However, the whole argument hinges on the *assumption* that biases are uniform within a subset. This assumption was also invoked for use with a pseudo-data simulation in an effort to justify the scoring procedure. Unless a subset consists of a single trial there is no way to justify this assumption. For instance, an agent-percipient pair might have some summer trials and some winter trials, with different biases. May et al. (1990, p. 202) also provide an illustration of how biases can vary.

Discussion

The problems enumerated above are severe. For a number of reasons it is most surprising that the research departs from usually expected criteria. Jahn and Dunne seem to have been aware of criticisms of earlier work for they say:

> Irrespective of the particulars, this type of experiment has become popular in contemporary parapsychological research because of its relatively high yield, low cost, and evident range of possible practical applications. It has also become a target of much critical commentary because of its high vulnerability to sloppy protocols and data-processing techniques. The most frequently cited criticisms include possibilities of inadvertent or deliberate sensory cuing of either participant, inadequately randomized target selection practices, shared conscious or subconscious preferences of the percipient and agent, and faulty statistical treatment. (Jahn & Dunne, 1987, p. 157)

Ironically, the above-mentioned issues have not been adequately recognized by PEAR in regard to their own program. Remote-viewing research has been conducted for over 15 years and is thus not a recently developed procedure in parapsychology. In new areas of investigation in any science, one expects some deficiencies, but one also expects a steady improvement of methods as the studies continue. The procedures and methodological issues of remote viewing have received considerable attention in the literature. The difficulties are easy and inexpensive to overcome with correct experimental design.

Some might argue that if one compares the "flawed" segments with the "unflawed" and no statistically significant difference is found, then one need not worry about the flaws. Even if such a comparison were contemplated, it would require a valid statistical baseline. It is not clear whether any of the trials involved adequate randomization (details are lacking in the reports), so we are unable to make legitimate comparisons. Even if we could make such comparisons, the lack of a statistically significant difference would not imply that a difference didn't exist. It could mean that the sample sizes were too small to provide enough power to detect a difference; the number of "unflawed" trials is likely to be small, and maybe even zero. (For a discussion of statistical power and sample size in parapsychology see Utts, 1988.)

The deficiencies provide plausible mechanisms for the significant findings. While the actual effects of the problems we have outlined are too complex to fully evaluate, it is instructive to see how much of an influence would be necessary to account for the observed results.

The mean score for the universal mismatch distribution was .5025, and for the actual formal trials it was .5447. How many additional descriptors would have to be answered correctly, on average, to account for the difference of .0422? Here we will make the invalid assumption that the fundamental statistical procedure of PEAR is appropriate. PEAR's descriptor weights are unknown to us, but can probably be realistically approximated by deriving values from Table C-1 (Dunne et al., 1983, p. 122). It can be easily shown that the mean value of the denominator of the score is approximately 60. Therefore, the difference of .0422 between mismatch and match scores represents an approximate increase of 2.532 in the numerator. The average contribution per correctly answered question is 2.00. Therefore, slightly better than a one question average advantage per trial, due to inflation of the matched scores or deflation of the mismatched scores, or a combination, would account for the difference.

Conclusions

The PEAR remote-viewing experiments depart from commonly accepted criteria for formal research in science. In fact, they are undoubtedly some of the poorest quality ESP experiments published in many years. The defects provide plausible alternative explanations. There do not appear to be any methods available for proper statistical evaluation of these experiments because of the way in which they were conducted.

Appendix

Consequences of an Inappropriate Statistical Model

Dunne et al. (1989, p. 13) specifically noted that targets were selected *without replacement*. This has important consequences for statistical analysis. It has been recognized for a long time that the analysis must consider the dependence caused by the without-replacement condition. In fact, Puthoff, Targ, and May (1981) even published a reanalysis of some of the early SRI remote-viewing experiments so that this dependence was taken into account. Other examples of appropriate analyses can be seen in Markwick (1988) and Schlitz and Gruber (1980).

An additional problem arises because in constructing their baseline distribution, PEAR did not include the diagonal elements of the matrix (i.e., the scores resulting from the correctly matched targets and responses). The failure to include these elements in the baseline distribution can artifactually enhance significance levels, as we will show.

TABLE 1
DESCRIPTORS AND PROBABILITIES

Descriptor number	A priori probability	Optimal response	Probability of being correct
1	.397	N	.603
2	.215	N	.785
3	.486	N	.514
4	.589	Y	.589
5	.154	N	.846
6	.229	N	.771
7	.537	Y	.537
8	.458	N	.542
9	.621	Y	.621
10	.715	Y	.715
11	.710	Y	.710
12	.248	N	.752
13	.439	N	.561
14	.154	N	.846
15	.360	N	.640
16	.290	N	.710
17	.743	Y	.743
18	.650	Y	.650
19	.547	Y	.547
20	.706	Y	.706
21	.495	N	.505
27	.421	N	.579
28	.612	Y	.612
29	.257	N	.743
30	.645	Y	.645
			19.840

Expected number of correct bits using the optimal-guessing strategy:

Score = 19.840/30 = .6613

The consequences of these two problems can be illustrated by examining the ratings matrix of Schlitz and Gruber (1980, p. 311). Their correctly computed probability value, by a direct count of permutations, was 4.7×10^{-6}. We can implement the PEAR method and form a distribution of the off-diagonal elements. The mean of that distribution is 198.10, and the standard deviation is 92.21. A z score can be calculated for each of the 10 diagonal elements. The resulting composite z score is 6.410, which corresponds to a p value of 7.28×10^{-11}.

We grant that the distribution is not normal, but it can be shown that the artifactual inflation of significance is not attributable to nonnormality. Fortunately we can compute an exact probability value for a virtually equivalent statistical model (making the same false assumption about appropriateness of the baseline distribution). The sum of the 10 largest off-diagonal elements is less than the sum of the diagonal elements. Conservatively, this results in a probability of $p = 80!10!/90! = 1.748 \times 10^{-13}$; the corresponding normal approximation gives $p = 5.35 \times 10^{-12}$.

As can be seen, in this case the PEAR method gives a probability value incorrect by about five orders of magnitude. The statistical method of PEAR is fundamentally flawed.

Optimal Guessing

In an article on deception, Hansen (1990, p. 52) mentioned that an optimal-guessing strategy may allow a sophisticated form of cheating. Even without conscious deception, some subjects might have a response pattern which would naturally produce high scores. We will describe here how such a strategy might be implemented. As mentioned earlier, PEAR used five different formulas to calculate scores. We will illustrate the optimal-guessing strategy with PEAR's susceptible Method A. Tables 41 and 42 of Dunne et al. (1983) show that Method A tended to produce more extreme p values than the other methods.

With Method A (Dunne et al., 1983, p. 12), a target descriptor list is compared with that of the response, and the number of corresponding matching bits is tallied; that is, if both lists give a positive (or negative) response for a descriptor, the score is incremented by one; if they do not match, the score is not increased. The resulting number is divided by 30 (the total number of bits). Thus a score of .5 would be associated with 15 bits that matched. Sixteen matching bits produce a score of .533.

As described earlier, in order to establish a baseline, Dunne et al. take targets and responses and pair all targets with mismatched responses. For each pairing of a target and response, a score is generated. These scores form a distribution. The mean and standard deviation are computed. For any numeric score, a z score can be directly calculated as previously described.

The optimal-guessing strategy consists simply of saying yes to descriptors with an a priori probability above .5 and no to descriptors which have an a priori probability below .5. Table C-1 on page 122 of Dunne et al. (1983) gives the probability of each descriptor's occurring for a target.

The expected score shown in Table 1 corresponds to a z score of 1.047 as computed by the method used by Dunne et al. (1983): $z = (.6613 - .5510)/.1053$ where the parameters are given on their page 92.

Using the optimal-guessing strategy, the expected z score for one trial is 1.047. Thus, for an experiment with 100 trials, the expected z score would be 10.47. Please note that the above analysis is an approximation that assumes independence among items on the descriptor list.

References

ALCOCK, J. E. (1990). *Science and supernature: A critical appraisal of parapsychology.* Buffalo, NY: Prometheus Books.

ALEXANDER, J. B., GROLLER, R., & MORRIS, J. (1990). *The warrior's edge.* New York: William Morrow.

BELOFF, J. (1989). Review of *Margins of reality: The role of consciousness in the physical world* by Robert G. Jahn and Brenda J. Dunne. *Journal of the Society for Psychical Research,* **55,** 364–368.

BRAUDE, S. E. (1987). How to dismiss evidence without really trying. *Behavioral and Brain Sciences,* **10,** 573–574.

CHILD, I. L. (1978, September–October). Some educational uses of the Maimonides slides. *Parapsychology Review,* **9**(5), 24–25.

DR. STANFORD replies . . . (1986, November–December). *Parapsychology Review,* **17**(6), 14–15.

DRUCKMAN, D., & SWETS, J. A. (Eds.). (1988). *Enhancing human performance: Issues, theories, and techniques.* Washington, DC: National Academy Press.

DUNNE, B. J., DOBYNS, Y. H., & INTNER, S. J. (1989). *Precognitive remote perception III: Complete binary data base with analytical refinements* (Technical Note PEAR 89002). Princeton, NJ: Princeton Engineering Anomalies Research, School of Engineering and Applied Science, Princeton University.

DUNNE, B. J., JAHN, R. G., & NELSON, R. D. (1983). *Precognitive remote perception* (Technical Note PEAR 83003). Princeton, NJ: Engineering Anomalies Research Laboratory, School of Engineering/Applied Science, Princeton University.

EDGE, H. L., MORRIS, R. L., RUSH, J. H., & PALMER, J. (1986). *Foundations of parapsychology: Exploring the boundaries of human capability.* Boston: Routledge & Kegan Paul.

ELLISON, A. (1988). *The paranormal: A scientific exploration of the supernatural.* New York: Dodd, Mead.

GOODFELLOW, L. D. (1938). A psychological interpretation of the results of the Zenith radio experiments in telepathy. *Journal of Experimental Psychology,* **23,** 601–632.

HACKING, I. (1988). Telepathy: Origins of randomization in experimental design. *Isis,* **79,** 427–451.

HANSEL, C. E. M. (1989). *The search for psychic power: ESP & parapsychology revisited.* Buffalo, NY: Prometheus Books.

HANSEN, G. P. (1990). Deception by subjects in psi research. *Journal of the American Society for Psychical Research,* **84,** 25–80.

HOUCK, J. (1986). Associative remote viewing. *Archaeus,* **4,** 31–37.

HUMPHREY, B. S., MAY, E. C., & UTTS, J. M. (1988). Fuzzy set technology in the analysis of remote viewing. *Proceedings of presented papers: The 31st annual convention of the Parapsychological Association,* pp. 378–394.

HYMAN, R. (1977, May–June). Psychics and scientists: "Mind-reach" and remote viewing [Review of *Mind-Reach* by Russell Targ & Harold Puthoff]. *The Humanist,* pp. 16–20.

HYMAN, R. (1988). Psi experiments: Do the best parapsychological experiments justify the claims for psi? *Experientia,* **44,** 315–322.

HYMAN, R. (1989). *The elusive quarry: A scientific appraisal of psychical research.* Buffalo, NY: Prometheus Books.

HYMAN, R., & HONORTON, C. (1986). A joint communiqué: The psi ganzfeld controversy. *Journal of Parapsychology,* **50,** 351–364.

JAHN, R. G. (1982). The persistent paradox of psychic phenomena: An engineering perspective. *Proceedings of the IEEE,* **70**(2), 136–170.

JAHN, R. G., & DUNNE, B. J. (1986). On the quantum mechanics of consciousness, with application to anomalous phenomena. *Foundations of Physics,* **16,** 721–772.

JAHN, R. G., & DUNNE, B. J. (1987). *Margins of reality: The role of consciousness in the physical world.* New York: Harcourt Brace Jovanovich.

JAHN, R. G., DUNNE, B. J., & JAHN, E. G. (1980). Analytical judging procedure for remote perception experiments. *Journal of Parapsychology,* **44,** 207–231.

JAHN, R. G., DUNNE, B. J., & JAHN, E. G. (1981). Analytical judging procedure for remote perception experiments. In W. G. Roll, J. Beloff (Eds.), and J. McAllister (Assistant Ed.), *Research in parapsychology 1980* (pp. 21–23). Metuchen, NJ: Scarecrow. (Abstract).

JAHN, R. G., DUNNE, B. J., & NELSON, R. D. (1987). Engineering anomalies research. *Journal of Scientific Exploration,* **1,** 21–50.

JAHN, R. G., DUNNE, B. J., NELSON, R. D., JAHN, E. G., CURTIS, T. A., & COOK, I. A. (1983). Analytical judging procedure for remote perception experiments: II. Ternary coding and generalized descriptors. In W. G. Roll, J. Beloff, & R. A. White (Eds.), *Research in parapsychology 1982* (pp. 159–163). Metuchen, NJ: Scarecrow. (Abstract)

KENNEDY, J. E. (1979a). Methodological problems in free-response ESP experiments. *Journal of the American Society for Psychical Research,* **73,** 1–15.

KENNEDY, J. E. (1979b). More on methodological issues in free-response psi experiments. *Journal of the American Society for Psychical Research,* **73,** 395–401.

MARKWICK, B. (1988). Re-analysis of some free-response data. *Journal of the Society for Psychical Research,* **55,** 220–222.

May, E. C., Utts, J. M., Humphrey, B. S., Luke, W. L. W., Frivold, T. J., & Trask, V. V. (1990). Advances in remote-viewing analysis. *Journal of Parapsychology*, **54,** 193–228.

McConnell, R. A. (1987). *Parapsychology in retrospect: My search for the unicorn.* Pittsburgh, PA: Author.

Morris, R. L. (1978). A survey of methods and issues in ESP research. In Stanley Krippner (Ed.), *Advances in parapsychological research 2: Extrasensory perception* (pp. 7–58). New York: Plenum.

Morris, R. L. (1982). An updated survey of methods and issues in ESP research. In Stanley Krippner (Ed.), *Advances in parapsychological research 3* (pp. 5–40). New York: Plenum.

Nash, C. B. (1986). *Parapsychology: The science of psiology.* Springfield, IL: Charles C Thomas.

Nelson, R. D., & Dunne, B. J. (1987). Attempted correlation of engineering anomalies with global geomagnetic activity. In D. H. Weiner & R. D. Nelson (Eds.), *Research in parapsychology 1986* (pp. 82–85). Metuchen, NJ: Scarecrow. (Abstract)

Nelson, R. D., Jahn, R. G., & Dunne, B. J. (1986). Operator-related anomalies in physical systems and information processes. *Journal of the Society for Psychical Research*, **53,** 261–285.

Nelson, R. D., & Radin, D. I. (1987). When immovable objections meet irresistible evidence: A case of selective reporting. *Behavioral and Brain Sciences*, **10,** 600–601.

Palmer, J. (1985). *An evaluative report on the current status of parapsychology* (Contract Number DAJA 45-84-M-0405). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Puthoff, H. E., Targ, R., & May, E. C. (1981). Experimental psi research: Implications for physics. In R. G. Jahn (Ed.), *The role of consciousness in the physical world* (pp. 37–86). Boulder, CO: Westview.

Rhine, J. B., & Pratt, J. G. (1962). *Parapsychology: Frontier science of the mind.* Springfield, IL: Charles C Thomas. (Originally published 1957)

Rockwell, T., & Rockwell, W. T. (1988). Review of *Margins of reality: The role of consciousness in the physical world* by Robert G. Jahn and Brenda J. Dunne. *Journal of the American Society for Psychical Research*, **82,** 359–365.

Rockwell, T., & Rockwell, W. T. (1989, March–April). Review of *Margins of reality: The role of consciousness in the physical world* by Robert G. Jahn and Brenda J. Dunne. *New Realities*, **9** (4), 65–66.

Schlitz, M., & Gruber, E. (1980). Transcontinental remote viewing. *Journal of Parapsychology*, **44,** 305–317.

Schlitz, M., & Gruber, E. (1981). Transcontinental remote viewing: A rejudging. *Journal of Parapsychology*, **45,** 233–237.

Scott, C. (1988). Remote viewing. *Experientia*, **44,** 322–326.

Stokes, D. M. (1978a). Review of *Research in parapsychology, 1976 (Abstracts and papers from the nineteenth annual convention of the Parapsychological Association). Journal of Parapsychology*, **42,** 70–76.

Stokes, D. M. (1978b). Review of *Research in parapsychology, 1977 (Abstracts and papers from the twentieth annual convention of the Parapsychological Association). Journal of Parapsychology*, **42,** 313–319.

Stokes, D. M. (1981). Review of *Research in parapsychology, 1979 (Abstracts and papers from the twenty-second annual convention of the Parapsychological Association). Journal of Parapsychology*, **45,** 174–183.

Thouless, R. H., & Brier, R. M. (1970). The stacking effect and methods of correcting for it. *Journal of Parapsychology*, **34,** 124–128.

Utts, J. (1988). Successful replication versus statistical significance. *Journal of Parapsychology*, **52,** 305–320.

Utts, J. (in press). Analyzing free-response data—A progress report.

Princeton Arms North 1, Apt. 59
Cranbury, NJ 08512

Division of Statistics
University of California
Davis, CA 95616

5 Thorncroft
Hornchurch
Essex RM11 1EU
England