Information and Uncertainty in Remote Perception Research

Brenda J. Dunne and Robert G. Jahn

Princeton Engineering Anomalies Research
Princeton University
Princeton NJ 08544-5263
e-mail: pearlab@princeton.edu

Abstract—This article has four purposes: 1) to present for the first time in archival form all results of some 25 years of remote perception research at this laboratory; 2) to describe all of the analytical scoring methods developed over the course of this program to quantify the amount of anomalous information acquired in the experiments; 3) to display a remarkable anti-correlation between the objective specificity of those methods and the anomalous yield of the experiments; and 4) to discuss the phenomenological and pragmatic implications of this complementarity. The formal database comprises 653 experimental trials performed over several phases of investigation. The scoring methods involve various arrays of descriptor queries that can be addressed to both the physical targets and the percipients' description thereof, the responses to which provide the basis for numerical evaluation and statistical assessment of the degree of anomalous information acquired. Twenty-four such recipes have been employed, with queries posed in binary, ternary, quaternary, and ten-level distributive formats. Thus treated, the database yields a composite z-score against chance of 5.418 ($p = 3 \times 10^{-8}$, one-tailed).

   Numerous subsidiary analyses agree that these overall results are not significantly affected by any of the secondary protocol parameters tested, or by variations in descriptor effectiveness, possible participant response biases, target distance from the percipient, or time interval between perception effort and agent target visitation. However, over the course of the program there has been a striking diminution of the anomalous yield that appears to be associated with the participants' growing attention to, and dependence upon, the progressively more detailed descriptor formats and with the corresponding reduction in the content of the accompanying free-response transcripts. The possibility that increased emphasis on objective quantification of the phenomenon somehow may have inhibited its inherently subjective expression is explored in several contexts, ranging from contemporary signal processing technologies to ancient divination traditions. An intrinsic complementarity is suggested between the analytical and intuitive aspects of the remote perception process that, like its more familiar counterpart in quantum science, brings with it an inescapable uncertainty that limits the extent to which such anomalous effects can be simultaneously produced and evaluated.

Keywords: remote perception—remote viewing—anomalous information acquisition—consciousness-related anomalies—uncertainty—complementarity—PEAR—engineering anomalies—analytical judging

Man also possesses a power by which he may see his friends and the circumstances by which they are surrounded, although such persons may be a thousand miles away from him at that time.

Paracelsus

I. Introduction and Background

This concise statement of the remote perception hypothesis was proffered by the renowned 16th-century physician and philosopher, Paracelsus, in a section of his writings devoted to the role of "active imagination" in man's representation of his universe.[1] His observation was certainly not the first recorded allusion to such anomalous human capabilities. This "power" has been acknowledged in virtually every culture since the dawn of human civilization, and invoked under a multitude of names including, among many others, divination, prophecy, oracle, scrying, clairvoyance, and second sight.

In the more recent history of Western science, a considerable body of literature describing scholarly investigations of "extrasensory perception" already had been amassed when, in the mid-1970s, Puthoff and Targ at Stanford Research Institute introduced a new scientific protocol for empirical investigation of the phenomenon they termed "remote viewing."[2,3] Their procedure required one individual, referred to as the "percipient," to attempt to describe the geographical ambience surrounding another person, the "agent," whose location was inaccessible to the percipient by any known sensory means. Their striking data included many perceptions that were virtually photographic in accuracy, and produced an overall statistical yield well beyond chance expectations. Over the subsequent quarter century, numerous replications of the original SRI studies have been reported,[4–16] including a number of originally classified government-sponsored investigations,[17–21] most of which display the ambiguous mixtures of successes and failures that seem to characterize most serious anomalies research. Notwithstanding, the majority of these studies demonstrate a sufficient degree of anomalous information acquisition to justify continued scholarly exploration of this mystifying process.

One of the largest extant databases, comprising 653 formal and 126 non-formal experimental trials, was produced between 1976 and 1999 as one of the three major components of the Princeton Engineering Anomalies Research (PEAR) program. The other two segments, concerning anomalies in human/machine interactions and theoretical modeling, have been reported extensively in this journal and elsewhere. The purpose of this paper is to describe the procedures and summarize the full results of our remote perception studies, and to explore their implications for better comprehension of this currently inexplicable communication capability. To achieve this most concisely, we shall refer frequently to a number of earlier publications and technical reports wherein all the datasets and analytical methods are presented in greater detail.[22–26]

The first phase of this PEAR work evolved from a body of prior experiments conducted between 1976 and 1979 by one of the authors (B.J.D.) at Mundelein

College in Chicago and subsequently at the University of Chicago,[27,28] which utilized human-judge ranking procedures similar to those of the earlier SRI studies.[29] Despite the impressive yield of these experiments, concerns regarding evident vagaries and possible subjective biases in the judges' interpretations, or even anomalous inputs on their part, predicated a more quantitative approach to data evaluation.[30] A primary focus of the subsequent PEAR studies has been on the development of analytical judging procedures capable of rendering the free-response raw data into forms amenable to more rigorous quantification and analysis. Beyond the acquisition and analysis of large composite databases, a number of secondary experimental variables, such as the effect of multiple percipients, alternative target selection procedures, and the dependence of the phenomenon on spatial and temporal separations, have also been explored. Inspired by a section of Puthoff and Targ's 1976 paper[3] wherein they alluded to the ability of some of their percipients to describe target scenes even before the target had been identified, much less visited, the majority of the PEAR trials have been acquired in this precognitive mode. And since many of the percipients maintain that their experiences are not, strictly speaking, of a simple visual nature, the term ''precognitive remote perception,'' or PRP, has been preferred.

II. Protocol

In its basic form, the PEAR protocol requires a percipient to describe an unknown remote geographical target where an agent is, was, or will be situated at a prescribed time. The target location is selected randomly before each trial from a large pool of potential targets, prepared previously by an individual not otherwise involved in the experiment. The contents of this pool are stored in separate sealed envelopes, randomly numbered, and maintained so that no agent or percipient has access to them. Prior to a given trial, the target is designated by generation of a random number that identifies one of the envelopes, which then is delivered, still sealed, to the agent, who opens it and follows instructions to locate the target. This "instructed" mode of target selection is complemented by a "volitional" protocol option, typically followed when the agent is traveling on an itinerary unknown to the percipient, in a region for which no prepared pool exists. In these trials, the agent simply selects the target from among the various local sites accessible at the time specified for the trial.

In either version, the percipient is asked to spend 15 to 20 minutes attempting to visualize or experience the target and to record these impressions in a free-response, stream-of-consciousness form, either orally into a tape recorder or in writing, optionally including drawings. Unlike some of the procedures followed at SRI and elsewhere, where percipients are trained to use particular strategies, or where perceptions are generated in a laboratory setting with an experimenter present and actively eliciting information, PEAR percipients are free to choose their own subjective strategies and physical locations, and experimenters are not

present during the perception process. While the majority of data have been acquired in the precognitive mode, wherein the perceptions are generated and recorded before the target is selected, a substantial subset of trials have been executed in a retrocognitive mode, wherein perceptions are generated after the agent has visited the target, and a smaller number have been performed in "real time." In all cases, strict precautions are taken to ensure that perceptions are recorded and filed before percipients have any sensory access to information about the targets, and no ordinary means of communication between percipients and agents is available until after that point.

The agents, who in almost all cases are known to the percipients, are asked to situate themselves at the target sites at the agreed-upon times and to immerse themselves in the scenes for about 15 minutes. At the close of the visitation periods, they record their impressions of the target scenes, supplementing them with hand-drawn sketches if desired, and whenever possible by one or more photographs to corroborate their verbal descriptions. Like the percipients, agents are free to employ their own subjective strategies. They simply are encouraged to attempt in some way to share their target experiences with the percipients.

All of the participants in the PEAR experiments have been uncompensated volunteers, none of whom has claimed exceptional abilities in this regard. No explicit tactical instructions are given, although an attitude of playfulness is encouraged and emphasis is placed on enjoyment of the experience, rather than on achievement *per se.* Transcript styles of individual percipients vary widely, ranging from a few cryptic details at one extreme, to lengthy impressionistic flows of imagery on the other. No systematic records have been maintained on the relative effectiveness of the various personal strategies deployed by the participants, or on any of their psychological or physiological characteristics. They are encouraged, however, to furnish subjective reports of their experiences, and these anecdotal descriptions have provided valuable glimpses into some of the more qualitative aspects of the underlying process. For example, several percipients have commented that they found it helpful to clear their minds, visualize a blank screen, and wait for an image of the agent to appear. Some agents report that they imagine that the percipients are with them at the target scene and that they carry on mental conversations with them, pointing out various aspects of the sites. On some occasions, agents have observed that they found their attention drawn to components of the scene that they had overlooked initially, only to discover later that these features had been part of the percipient's descriptions, almost as if the percipient's consciousness had guided their attention. Many participants have indicated that they feel more like they are sharing a common experience, rather than "transmitting" information from one person to another.

III. Analytical Judging Methods: Development and Initial Applications

As mentioned earlier, evaluation of the original Chicago experiments that had produced highly significant statistical results had been based on rankings

assigned by independent human judges to each of the free-response perceptions when compared with photographs of all the targets in its local series.[30] To assess the potential statistical impact of inter-judge variability in those studies, 27 transcripts comprising the first three experimental series had been subjected to repeated re-judging by five separate individuals. Although approximately half of these trials demonstrated a strong consistency in the ranks assigned by both the primary and secondary judges and confirmed the acquisition of significant extra-chance information, the others received a wide range of ranks, suggesting that the matches originally assigned to these trials had most likely been arbitrary. Also evident in this review was the inherent inefficiency of an approach whereby the entire informational content of a given perception was reduced to a single datum, ordinal at best, in a small experimental series.

Beyond the accumulation of new empirical data, the first major thrust of the embryonic PEAR program was an attempt to alleviate some of these shortcomings by developing standardized methods of quantifying the information content of the free-response data via a series of computer algorithms. The first step in this direction was the establishment of a code, or alphabet, of 30 simple binary descriptive queries that could be addressed to all targets and perceptions. The questions ranged broadly from factual, *e.g.* whether the scene was indoors or outdoors, whether water was present, etc., to more impressionistic, *e.g.* whether the scene was confined or expansive, noisy or quiet, etc. The responses, entered into a computerized database manager as strings of 30 bits, were submitted to an assortment of analytical scoring algorithms that could provide numerical evaluation of the thus-specified information content of any given trial, and once scored, the statistical merit of the perception results could be evaluated by an assortment of computerized analytical ranking procedures.[22] Specifically, the algorithms scored each transcript against all the targets in the pool and then ranked them in order of descending score.

While still dependent upon a ranking procedure, this descriptor-based process had the advantages that such ranking could proceed on a more standardized analytical basis and that many more alternative targets could be ranked by the computer than by a human judge. As a first test of this approach, one series of eight trials from the earlier Chicago database was encoded *ex post facto* into the binary format by five independent encoders. Reassuringly, most of the responses were found to be in close agreement with each other, *i.e.*, the computer-assigned ranks of the better trials were highly consistent with those of the original human judges, and those of the weaker trials were comparably equivocal.

With these scoring methods so qualified, 35 new trials were generated following the same protocol used in the earlier experiments, but now the targets and perceptions were descriptor-encoded *ab initio* by the agents at the target sites and by the percipients after completing their free-response descriptions. Although the statistical results of these new trials were not as strong as those of the *ex post facto*–encoded data, they were still highly significant. Perhaps even more importantly, the general agreement among the various scoring algorithms

confirmed that the analytical methodology was indeed capable of providing reliable quantification of the intrinsically impressionistic remote perception data. To obviate the possibility that the particular list of descriptors employed somehow could process even random inputs to apparently significant scores, a "calibration" exercise was undertaken wherein artificial "target" and "perception" data matrices of the same size as the actual data matrices were constructed from the output of a random event generator. The same computational schemes were applied to various combinations of these, both with each other and with the true data, with results that were all well within chance expectation.[25]

With growing confidence in the viability of this analytical methodology, an additional 51 prior trials from Chicago and PEAR were then transcribed into the new descriptor format, increasing the total number of *ex post facto*–encoded trials to 59, comprising all the original human-judged trials that met formal protocol criteria and had adequate target documentation to permit such retrospective encoding. Here and henceforth, *formal* trials are defined as those that follow the standard protocol described earlier and also meet all of the following criteria:

1. The agent and percipient are specified to one another.
2. The date and time of the agent's target visitation are specified to the percipient.
3. The agent is present at the target within 15 minutes of the specified time and is consciously committed to his or her experimental role during that period.
4. Both agent and percipient produce verbal descriptions and complete the descriptor response forms.
5. Both agent and percipient have adequate familiarity with the application and interpretation of the descriptor questions and with the general protocol.
6. Photographs, written descriptions, or other substantiating target information are available.

By 1983, the 59-trial *ex post facto*–encoded database had been supplemented by 168 new *ab initio*–encoded trials, plus 73 others that for various reasons did not meet formal protocol criteria, bringing the total to 300. Of the non-formal trials, 21 were categorized as "questionable," where failure to meet the formal criteria was due to protocol violations, such as the lack of adequate substantiating target information, evidence that one or both of the participants did not understand the application or interpretation of the descriptor questions, or the vulnerability of the trial to sensory cueing. Another 52 trials were designated in advance as "exploratory," wherein intentional deviations from formal protocol, such as deliberately not informing the percipient of the agent's identity, or not specifying the time of target visitation, were undertaken.[24]

IV. Statistical Evaluations via Empirical Chance Distributions

Beyond its evident success in dispassionate ranking of the trials in any given experimental series, the descriptor-based scoring method offered a far more

desirable and powerful capability, *i.e.*, the direct calculation of the statistical merit of individual trial scores or groups of scores. To achieve this, an empirical "chance" distribution was constructed by scoring every perception in the 300-trial database against every possible target except its correct one, thus compounding a large array of deliberately mismatched scores, the distribution of which displayed classical Gaussian features and could serve as a statistical reference. Several variations of this scoring technique were explored, all of which consisted of calculating a score for each trial based on the proportion of matches and mismatches in the percipient and agent responses to the 30 descriptor queries, using a set of generalized *a priori* probabilities derived from the 300 targets comprising the database as descriptor weighting factors. For example, since more targets tended to be outdoors than indoors, a correct positive response to the query "Is the scene indoors?" was assigned a greater weight than a correct negative response, and its incremental contribution to the total score was proportionately larger. The sum of the score increments from all 30 descriptors constituted the "absolute score" for a given trial, which was then divided by some normalizing factor, such as the maximum score that would have been achieved had all 30 target and perception descriptor responses agreed, yielding a "normalized score." The statistical merit of this normalized score was then established by comparing it with the chance distribution of similarly normalized mismatched scores.

The descriptor response check sheets also contained a column labeled "unsure" in addition to the standard "yes" and "no" options, which permitted participants to indicate any ambiguities they might experience in relating their subjective impressions in strictly binary terms. These "unsure" responses were disregarded in the binary calculations, but they provided the basis for investigating the potential benefits of ternary-based algorithms.[23] Seven such ternary scoring methods were explored, all of which showed good internal consistency, but none of which indicated any substantial advantage over the binary calculations. Given their added computational complexity, subsequent study was limited to only five binary-based methods:

- Method A: The number of descriptors answered correctly, divided by the total number of descriptors (*i.e.*, a count of the numerical fraction of correct responses, ignoring the *a priori* descriptor probabilities).
- Method B: The sum of all descriptors answered correctly, each weighted by the reciprocal of its *a priori* probability, divided by the sum of all descriptors so weighted. (This method weighted the value of correct responses in inverse proportion to their *a priori* probabilities, and normalized the score by the highest possible score obtainable by this method for a given target.)
- Method C: The same numerator as Method B, divided by the total number of descriptors, normalized by the "chance" score derived from the *a priori* probabilities.
- Method D: The sum of all descriptors correctly answered "yes," each weighted by the reciprocal of its *a priori* probability, plus the unweighted

sum of all descriptors answered "no," the total divided by the sum of all descriptors labeled "yes" in the target, each weighted by the reciprocal of its *a priori* probability, plus the unweighted sum of all descriptors labeled "no" in the target, with the resultant score weighted by the highest possible score for that target. (This process effectively removed from the calculation those descriptors on which the percipient responded negatively, whether correctly or incorrectly, and thereby served to countervene use of a negative response to imply ignorance of the descriptor, rather than its explicit absence.)

- Method E: The same numerator as Method D, divided by the total number of descriptors, *i.e.*, by the "chance" score.

Table 1 summarizes the results of these 300 trials, grouped by experimental criteria, as assessed by each of these five recipes.

The most instructive feature of these results is the consistency of anomalous yield across these five diverse scoring schemes. Regardless of the algorithm employed, for all but the exploratory trials the composite results indicate highly significant increments of anomalous information in the matched scores that are not present in the mismatched score distributions constructed from the same raw data. Even the null results of the 52 exploratory trials are informative in their indication that the features violated in these excursions from the standard protocol, *i.e.*, the percipients' knowledge of the agent or of the time of target visitation, may be requisites to generation of the anomalous effect. Given the evident insensitivity of the results to the particular scoring strategy deployed, it was agreed that only one method would henceforth be used as the standard for evaluating future binary-encoded trials. Method B was selected for this purpose, since it treated positive and negative descriptor responses in a symmetrical and intrinsically normalized fashion.

These results made it clear that the new analytical methodology was capable of relatively objective, quantitative assessment of the inherently subjective remote perception phenomenon. Unlike the less efficient, labor-intensive human judging methods, it not only could calculate individual trial scores, but could provide robust indications of the statistical quality of large databases. On the other hand, the analytical judging process introduced certain imperfections of its own. For example, the forced "yes" or "no" responses were limited in their ability to capture the overall ambience or context of a scene, or nuances of subjective or symbolic information that might be detected by human judges. Furthermore, while restricting the extracted information to the 30 specified binary descriptors minimized the reporting task for the participants, it precluded utilization of other potentially relevant features in the transcripts, such as specific colors, textures, architectures, or any other details not covered by the questions. These shortcomings were partially offset by the continued requirement that percipients first generate free-response descriptions from which the descriptor responses were then derived, a procedure intended to retain the

TABLE 1
Summary of Binary PRP Data as of 1983

Scoring method	Chance mean	Chance S.D.	Mean score	Composite z-score	Probability (one-tailed)	# Trials $p < .05$*	% Trials $p < .05$*
Formal data ($N = 227$)							
A	0.5610	.1053	0.6113	7.197	3×10^{-13}	28 (4)	12% (2%)
B	0.5042	.1207	0.5590	6.833	4×10^{-12}	40 (6)	18% (3%)
C	1.0005	.2380	1.1101	6.941	2×10^{-12}	35 (5)	14% (2%)
D	0.6512	.0935	0.6926	6.672	1×10^{-11}	33 (6)	15% (3%)
E	1.0034	.1330	1.0676	7.277	2×10^{-13}	35 (4)	14% (2%)
Formal plus questionable data ($N = 248$)							
A	0.5610	.1053	0.6071	6.894	3×10^{-12}	30 (4)	12% (2%)
B	0.5042	.1207	0.5536	6.442	6×10^{-11}	42 (7)	17% (3%)
C	1.0005	.2380	1.0998	6.574	2×10^{-11}	37 (6)	15% (2%)
D	0.6512	.0935	0.6887	6.321	1×10^{-10}	34 (6)	14% (2%)
E	1.0034	.1330	1.0619	6.924	2×10^{-12}	37 (4)	15% (2%)
Exploratory data ($N = 52$)							
A	0.5610	.1053	0.5538	−0.493	(.31)	0 (3)	0% (6%)
B	0.5042	.1207	0.5023	−0.115	(.45)	2 (3)	4% (6%)
C	1.0005	.2380	1.0277	0.824	.20	3 (2)	6% (4%)
D	0.6512	.0935	0.6419	−0.719	(.24)	1 (2)	2% (4%)
E	1.0034	.1330	1.0246	1.148	.13	5 (1)	10% (2%)
All data ($N = 300$)							
A	0.5610	.1053	0.5979	6.070	6×10^{-10}	30 (7)	10% (2%)
B	0.5042	.1207	0.5447	5.809	3×10^{-9}	44 (10)	15% (3%)
C	1.0005	.2380	1.0873	6.320	1×10^{-10}	40 (8)	13% (3%)
D	0.6512	.0935	0.6806	5.447	3×10^{-8}	35 (8)	12% (3%)
E	1.0034	.1330	1.0554	6.773	6×10^{-12}	42 (5)	14% (2%)

Note: The original version of this table, published in Technical Report 83003, contained an error that inadvertently inflated the results from Method A, suggesting that this method produced larger effects than the others. With this corrected, the results are reasonably consistent across all five methods.
* Numbers in parentheses indicate number of trials with negative z-scores, $p < .05$.

spontaneity of the PRP experience as well as to preserve the raw data in a suitable format for further study. Nonetheless, it became evident that after several experiences with the descriptor utilization, many participants tended to limit their attention and descriptions to those features that they now knew were specific to the questions.

These limitations notwithstanding, the evident advantages of the analytical judging techniques encouraged further exploration, beginning with a comprehensive evaluation of the effectiveness of the individual descriptors in constructing the trial scores. From this it was determined that the entire group of descriptors, originally selected by some combination of anecdotal experience and intuition, actually comprised a reasonably uniform set in terms of their effectiveness in quantifying informational bits across a broad range of target types. None was found to be extremely effective; none was seriously deficient. Sub-division of the descriptors into classifications of natural *vs.* man-made,

objective *vs.* subjective, permanent *vs.* transient, and indoor *vs.* outdoor, also indicated no significant differences in effectiveness. The interdependence among the various descriptors, *e.g.* that outdoor scenes were less likely to be confined, or that indoor scenes were less likely to involve airplanes or road vehicles, was also explored by a variety of statistical methods, all of which confirmed that while such correlations might blunt the incisiveness of the full descriptor net somewhat, they could not compromise the validity of the results.[24,25,31]

Thus, by the close of this phase of the program, a number of useful general conclusions had emerged:

1. Although the various methods produced differing scores for some of the individual trials, the overall statistical yield was uniformly highly significant and relatively insensitive to the particular scoring and normalizing recipes employed.
2. There was general agreement between the results of the various analytical methods and those of the impressionistic assessments by human judges, particularly for the perceptions of higher statistical merit.
3. The use of ternary descriptor responses, wherein participants were offered the option of "passing" on a given descriptor, did not yield sufficiently more consistent or accurate results compared to the binary methods to justify the added computational complexity.
4. Defining a "universal" target pool in terms of a sufficiently large number of actual targets made it possible to calculate a set of generalized *a priori* descriptor probabilities that could be used for scoring any individual perception efforts in the database, regardless of its particular local series pool.
5. Calculation of the statistical merit of individual perception efforts by reference to an empirical chance distribution, derived from a large number of deliberately mismatched targets and perceptions, proved to be a far more powerful strategy than the computerized analytical ranking within individual small series.
6. The 30 descriptors, originally chosen through a combination of empiricism and intuition, although clearly non-independent, nonetheless displayed a reasonably flat profile of effectiveness in building the scores of the significant transcripts.

V. Secondary Parameters

With the effectiveness of the analytical methodology thus established and the computerized ranking procedures superseded by the more powerful statistical procedure that compared the scores of individual trials or groups of trials with a "universal" mismatch distribution, a second phase of *ab initio*–encoded data generation was initiated that extended over several years. Since the protocols,

descriptor questions, and scoring algorithms remained identical to those deployed in the previous phase, these new trials could legitimately be combined with the earlier data to provide a larger database for structural segmentations. By 1988 the total PEAR PRP binary-descriptor database consisted of 411 trials, produced by a total of 48 participants. Of these, 336 trials qualified as formal, 54 as exploratory, and 21 as questionable. Of the 336 formal trials, 125 followed the instructed protocol, wherein the target was selected at random from a pre-existing pool, and 211 utilized the volitional protocol, wherein the agent was in an area for which no prepared pool existed.

Sorting the data by another criterion, 291 trials, 216 of which qualified as formal, were generated under the standard protocol wherein a single percipient attempted to describe the location of a single agent. In the remaining 120 trials, all of which met the formal criteria, two or more percipients addressed the same target. The number of percipients addressing a given target ranged from two to seven, and each perception was scored as a separate trial against its appropriate target. In all but two of the multiple-percipient trials, the percipients were aware that others were involved in the experiment, although they did not always know their identities. The participating percipients always were separated spatially from each other and, in most cases, attempted their perception efforts at different times. One series of formal trials and a few of the exploratory trials involved more than one agent, but in each of these cases only one, pre-specified, set of target encodings was included in the scoring process; the second set was used only for informal comparison.

Table 2 presents the summary statistics obtained using binary Method B for this combined PRP database and its various subsets. The empirical chance distribution used as a reference was derived from all the formal trials in this same database, and comprised more than 100,000 mismatched scores. In addition to the subsets addressing planned variations of the protocol, *e.g. ab initio vs. ex post facto* encoding, single *vs.* multiple percipients, and instructed *vs.* volitional assignment of targets, summaries for *ad hoc* subdivisions of the database by seasonal and regional target groupings are also included. For each independently calculated subset, the table displays the number of trials; the mean score; the effect size (defined as the mean z-score of all the trials in the given subset) with associated 99% confidence intervals; the standard deviation of the trial z-score distribution (expectation $= 1$); and the composite z-score (calculated by multiplying the effect size by the square root of the number of trials in the subset) with its associated one-tailed probability against chance. The last three columns list the number of trials in each subset with $z > 1.645$ ($p < .05$) (numbers in parentheses indicate $z < -1.645$); the corresponding percentage of those significant trials; and the percentage of scores where $p < .50$ (greater than the chance mean score). Each group is scored using the local *a priori* descriptor probabilities associated with that subset, and except for the groups labeled "All Trials" and "Non-Formal Trials," the various subsets consist of formal trials only. All are calculated with reference to the universal

TABLE 2
Binary PRP Data Summaries (Scoring Method B)

Subset	# Trials	Mean score	Effect size	99% Confidence interval	S.D. z-score	Composite z-score	Probability (one-tailed)	# Trials $p < .05$*	% Trials $p < .05$*	% Trials $p < .50$
All trials	411	.5364	.279	±.135	1.060	5.647	8×10^{-9}	47 (12)	11% (3%)	59%
Formal trials	336	.5447	.347	±.152	1.083	6.355	1×10^{-10}	44 (8)	13% (2%)	62%
Non-formal trials	75	.4969	−.046	±.278	0.910	−0.399	.655	3 (4)	4% (5%)	44%
Ab initio	277	.5345	.263	±.161	1.033	4.378	6×10^{-6}	31 (5)	11% (2%)	59%
Ex post facto	59	.5942	.754	±.417	1.203	5.792	3×10^{-9}	14 (2)	24% (3%)	75%
Single percipient	216	.5489	.382	±.194	1.098	5.613	1×10^{-8}	34 (6)	16% (3%)	60%
Multiple percipient	120	.5404	.312	±.251	1.049	3.416	3×10^{-4}	12 (3)	10% (3%)	63%
Instructed targets	125	.5653	.516	±.267	1.140	5.771	4×10^{-9}	23 (5)	18% (4%)	65%
Volitional targets	211	.5322	.244	±.191	1.066	3.549	2×10^{-4}	25 (3)	12% (1%)	60%
Summer trials	244	.5466	.363	±.183	1.099	5.663	7×10^{-9}	35 (5)	14% (2%)	65%
Winter trials	92	.5407	.315	±.286	1.043	3.017	1×10^{-3}	13 (2)	14% (2%)	57%
Chicago targets	31	.6189	.957	±.587	1.189	5.330	5×10^{-8}	10 (1)	32% (3%)	81%
Princeton targets	106	.5504	.394	±.286	1.110	4.060	2×10^{-5}	14 (3)	13% (3%)	62%
Targets elsewhere	199	.5267	.199	±.194	1.051	2.810	2×10^{-3}	20 (3)	10% (2%)	58%

* Numbers in parentheses indicate number of trials with negative z-scores, $p < .05$.

chance distribution of mismatched scores (N $=$ 106,602, mean $=$ .5025, and standard deviation $=$ .1216).

The overall results of these analyses leave little doubt, by any criterion, that the PRP perceptions contain considerably more information about the designated targets than can be attributed to chance guessing. Although the superior results of the *ex post facto* trials relative to the *ab initio* trials are particularly striking, little difference is found between single- and multiple-percipient performances, and there is no evidence of seasonal dependencies. (In assessing these results, it is important to keep in mind that the statistical *z*-scores reflect both the average effect size and the number of trials in each subset. So, for example, although the single-percipient data produce a substantially larger *z*-score than the smaller multiple-percipient subset, their relative effect sizes are very close and the large confidence intervals indicate that the two groups are statistically indistinguishable. Similar remarks pertain to the seasonal discriminations.)

The substantial difference between the yields of the *ex post facto* and *ab initio* data raise some concern that the former, on which the descriptor questions and methodology initially had been based, could have introduced a spurious score inflation into the composite database. Therefore, these analyses were repeated using only the formal *ab initio* data. The composite results of these 277 trials, presented in Table 3, continue to display a robust overall effect and confirm that the bottom-line yield of the overall PRP database cannot be discounted on the basis of any such inflation. It is interesting to note, however, that in this somewhat more restricted dataset the difference between the instructed and volitional subsets is considerably smaller and only marginally significant, and the geographical distinction between Princeton targets and those elsewhere, once the *ex post facto* Chicago trials are excluded, becomes statistically non-significant.

The difference between the average effect sizes of the instructed and volitional trials is worth closer examination since these two subsets might have been expected to display disparities in their empirical *a priori* descriptor probability estimates. Given the less formal nature of the target selection process in the volitional trials, it was possible that the agent's knowledge of the percipient's personal preferences or target response patterns could have influenced the target selection and representation, thereby introducing an undue bias into the volitional trial scores. In the full database, summarized in Table 2, there was indeed a statistically significant difference between the results of these two subsets ($z =$ 2.41), but it was actually the *instructed* subset that produced the larger effect size. The formal *ab initio* data only (Table 3) still showed a larger effect in the instructed trials, although the difference here was considerably smaller ($z =$ 1.73). Thus, the concern that the target selection process employed in the volitional trials might have contributed to artificial enhancement of the results appeared to be unfounded. If anything, these comparisons suggested that the volitional target selection process may actually have had an inhibitory effect on the phenomenon, rather than imposing an advantage.

TABLE 3
Formal *Ab Initio* Data Summaries (Scoring Method B)

Subset	# Trials	Mean score	Effect size	99% Confidence interval	S.D. z-score	Composite z-score	Probability (one-tailed)	# Trials $p < .05$*	% Trials $p < .05$*	% Trials $p < .50$
All trials	277	.5345	.263	±.161	1.034	4.378	6×10^{-6}	31 (5)	11% (2%)	59%
Single percipient	194	.5370	.284	±.197	1.063	3.949	4×10^{-5}	24 (6)	12% (3%)	56%
Multiple percipient	83	.5321	.243	±.275	0.974	2.215	.013	5 (1)	6% (1%)	64%
Instructed targets	94	.5416	.322	±.296	1.115	3.122	9×10^{-4}	11 (5)	12% (5%)	61%
Volitional targets	183	.5308	.233	±.194	1.020	3.148	8×10^{-4}	21 (1)	11% (.05%)	60%
Summer trials	195	.5374	.287	±.195	1.058	4.013	3×10^{-5}	24 (4)	12% (2%)	62%
Winter trials	82	.5308	.233	±.285	1.002	2.107	.018	7 (2)	9% (2%)	56%
Princeton targets	106	.5504	.394	±.281	1.125	4.060	2×10^{-5}	14 (4)	13% (4%)	62%
Targets elsewhere	171	.5243	.180	±.197	1.000	2.348	9×10^{-3}	16 (1)	9% (.05%)	59%

Numbers in parentheses indicate number of trials with negative z-scores, $p < .05$.

* Numbers in parentheses indicate number of trials with negative z-scores, $p < .05$.
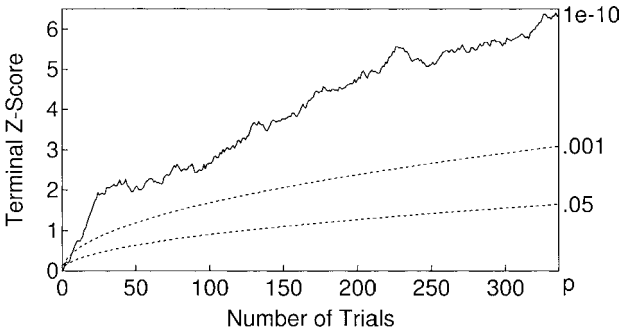
Fig. 1.   Cumulative deviation of 336 binary-encoded formal trials.

The magnitude and consistency of the anomalous yield in these data are presented graphically in Figure 1, where the results of all 336 formal trials are displayed in the form of a cumulative deviation of the actual scores from chance. Here, the stronger yield of the early *ex post facto* trials is strikingly evident. Nonetheless, the remainder of the trace, while less steep, also shows a clear and systematic deviation from chance expectation.

Further details on the analytical judging methodology and individual trial results, as well as examples of target photos and transcripts from some specific trials, may be found in Refs. 24–26, 32, and 33, and a process that verifies that the scores are not inflated by shared percipient/agent coding biases is described in Appendix A of this paper.

VI. Distance and Time Dependencies

Beyond the secondary parameters discussed in the previous section, a number of other variables were explored in the course of these experiments that proved helpful in illuminating some of the fundamental characteristics of the anomalous communication process. Two features of particular importance are the dependence of the results on the physical distance separating the percipient and the target, and on the time interval between the perception effort and the agent's visitation of the target. The spatial distances in this database ranged from less than one mile to several thousand miles, and the temporal separations from several days before to several days after target visitation. Figures 2 and 3 display the results of regression analyses of the dependence of the trial scores on these two parameters. In each, the horizontal dashed line denotes the empirical mean z-scores, the central dotted line indicates the linear regression fits to the data, and the outer dotted lines are the 95% confidence intervals thereof. Since the regressions are statistically indistinguishable from the lines of constant mean shift, we conclude that, within the ranges of this database, there are no significant correlations of effect size with either distance or time. In particular, when a regression of the data is plotted as a function of the reciprocal square of
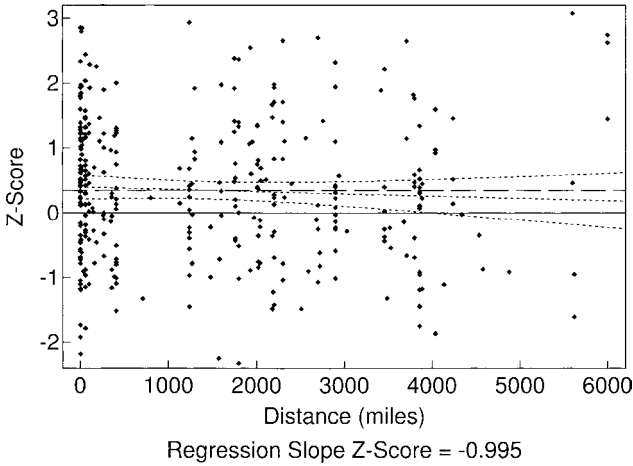
Fig. 2.   336 binary-encoded formal trial scores as a function of distance.

the distance, the results specifically refute any $1/r^2$ dependence of the anomalous "signal." Furthermore, if the data are segregated into subsets of the more extreme spatially and temporally displaced trials and those more proximate, the average effect sizes of the former remain statistically indistinguishable from those of the latter.[24,25]

The lack of evidence for attenuation of the remote perception yield with increased distance or time severely limits the possibilities for theoretical explication in terms of any known physical process. However, these findings did prompt the testable hypothesis that other anomalies being explored by PEAR might display similar non-local characteristics, and led to an extensive study of remote human/machine interactions. Here again, significant intention-correlated mean shifts have been observed that are statistically indistinguishable from those in the local experiments. Not only are the scales of these anomalous effects insensitive to intervening distance and time, but they display the same structural patterns as those of the corresponding local experiments.[34] Indeed, the similarities between the human/machine and remote perception results provided the first indications that these two forms of anomaly, previously regarded as distinct phenomena, actually might derive from the same mechanism of information exchange.

VII. FIDO Scoring

By 1985 the PEAR program had amassed a substantial body of experimental data that both confirmed the reality and robustness of the remote perception phenomenon and demonstrated the efficacy of the analytical scoring techniques. Although the *ab initio*–encoded trials had produced a smaller average effect size
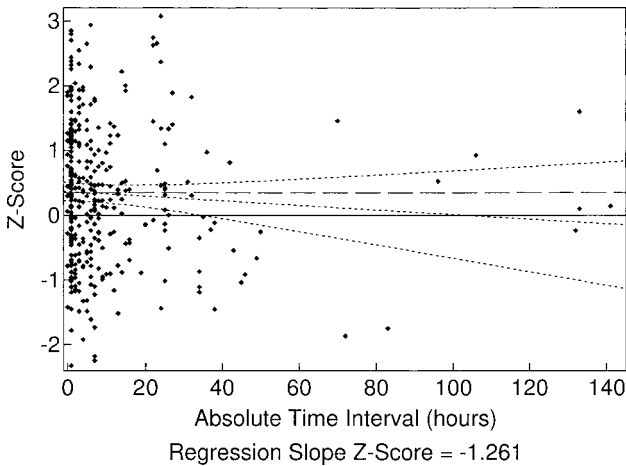
Fig. 3.   336 binary-encoded formal trial scores as a function of time.

than that of the *ex post facto* subset, this was attributed primarily to an inherent advantage for the earlier data of having the descriptor questions and analytical techniques based on those trials. The results of the *ab initio* experiments were still highly significant statistically, and the sacrifice of some of the impressionistic yield of the earlier efforts was deemed a reasonable price to pay for the capacity for more incisive quantitative measurement of the information content of the data. Notwithstanding, the diminished effect size prompted a new phase of investigation with the goal of achieving a better understanding of the cause of this attenuation and recovering the stronger yields obtained in the original experiments.

In the course of generating the *ab initio* data, several participants had complained that the forced binary responses seemed somewhat inhibitory and incapable of capturing many aspects of their experiences, suggesting that this might have contributed to the deterioration of the results. It was clearly evident that many of the target scenes, and most of the perceptions, contained ambiguous features that could not be answered easily with simple "yes" or "no" responses. For example, an agent might be indoors, but looking out a window at an outdoor scene, and thus unsure whether to characterize the scene as indoors or outdoors. A feature might have captured the agent's attention during the target visitation, but not have been an integral component of the scene itself, such as a brief conversational exchange with a passerby in an otherwise unpopulated area, complicating the response to the question "Are people present?" This problem was particularly evident in percipients' efforts to identify specific details from a perception that often emerged as a less than coherent stream of consciousness, much as in the difficulty of recalling features from fragments of dream imagery.

In an effort to make the analytical judging process more "user friendly," a quaternary descriptor response alternative was devised, playfully termed

FIDO, an acronym for "Feature Importance Discrimination Option." This new format provided participants with four response options for each descriptor: a rating of "4" identified a feature as a clearly dominant component of the scene; "3" meant the feature was present, but not particularly important; "2" indicated uncertainty as to the presence or absence of the feature; and "1" was a statement of the definite absence of the feature. Since implementation of the FIDO program required rewording of the descriptors, combination of the FIDO trials with the earlier databases was not feasible, but it did provide an opportunity to clarify or redefine some of the existing questions that had posed occasional interpretational difficulties. After an extensive assessment, which included having several people encode a variety of test scenes with the new quaternary descriptors and comparing their responses for consistency, a revised set of 32 descriptors was created and a new body of experiments undertaken. In all other respects, the same protocol was followed as in the earlier studies, although data were now generated on a trial-by-trial basis, rather than in series of arbitrary length. The FIDO program ran for four years, beginning in 1985, and produced a total of 167 trials.

The standard FIDO scoring matrix, illustrated below, assigned a score of 5 to each correctly matched response to options "absent" and "dominant," where there was agreement on the clear presence or absence of a given feature. A score of 4 was assigned to correct matches of "present" or "unsure." Mismatches of "absent" vs. "unsure," or "present" vs. "dominant," where percipient and agent agreed on the presence or absence of a feature but assigned it different degrees of importance, received a score of 3 if the percipient was less confident than the agent, but only 2 if the percipient was more confident. An "unsure" vs. "present" mismatch received a score of 2; mismatches of "absent" vs. "present," "or unsure" vs. "dominant," were assigned a score of 1; and a total mismatch of "dominant" vs. "absent" was scored as 0.

	Absent	Unsure	Present	Dominant	
Absent	5	3	1	0	↑
Unsure	2	4	2	1	
Present	1	2	4	2	Target
Dominant	0	1	3	5	↓

←   Perception   →

The scores derived from the 32 descriptor comparisons were added to produce a total score for each individual trial, as in the previous binary analyses. A matrix was then constructed that scored all the targets against all the perceptions, and the scores of the correct matches compared with the distribution of mismatched scores. Rather than attempting to establish *a priori* probabilities for these more complex descriptor options, the FIDO calculations were carried out using a method similar to binary Method A, which simply divided the sum of

the descriptor scores by the total number of descriptors, ignoring any *a priori* descriptor probabilities. The composite *z*-score thus calculated for the 167 FIDO trials was 1.735, indicating a marginally significant overall achievement, but one that was reduced even further from the high yield of the previous data.

Five alternative algorithms subsequently were applied *ex post facto* to these FIDO data in an effort to understand the cause of the lower yield and to devise more effective scoring strategies. Two of these methods simply returned the data to the original binary and ternary formats to ascertain whether the lower yield was attributable to an analytical insensitivity of the new technique or to poorer percipient performance. The binary reduction treated all responses of 4 or 3 as a ''yes,'' and all 2 or 1 responses as a ''no,'' while the ternary reduction treated a response of 4 as a ''yes,'' a response of 1 as a ''no,'' and a response of 2 or 3 as an ''unsure.'' A fourth method ignored everything but exact matches, assigning a score of 1 for each descriptor response in the perception that matched that in the target. Two additional methods allowed partial credit for close matches, similar to that of the standard FIDO algorithm. One assigned a score of 2 for an exact match and a score of 1 for an ambiguous match; the other assigned a weight of 4 to an exact match and a score of only 1 for an ambiguous match. A summary of the results produced by these six methods is presented in Table 4.

Other than the binary-reduction version, which produced nearly as many extra-chance ''misses'' as ''hits,'' the results from the other five methods all displayed relatively close concurrence, marginally significant composite *z*-scores, and effect sizes only about half that of the *ab initio* trials and only about a fifth as large as that of the *ex post facto* subset. Although the proportions of trials with positive scores were above 50% in all the calculations, neither these nor the numbers of significant trials exceeded chance expectation. Clearly, FIDO had not achieved its goal of enhancing the PRP yield, despite its potential sensitivity to subtle or ambiguous informational nuances in the data. Despite some variability among the *z*-scores calculated for individual trials by the different scoring methods, the general consistency across most of the scoring methods for the composite database suggested that the decreased yield was not directly due to inadequacies in the FIDO scoring algorithms, *per se*, but to a more generic suppression of the anomalous information channel.

This suspicion was reinforced by a supplemental exercise in which an independent human judge was asked to rank the fits between the agents' free-response transcripts and their coded descriptors. This ranking effort was admittedly subjective and arbitrary, and complicated by the varied lengths of transcripts and the presence or absence of drawings, photos, or other illustrative material. However, of the 167 targets, the judge determined that 162 (97%) showed reasonably good correspondences between the agents' verbal descriptions and their descriptor responses. A similar exercise was performed on the percipients' encodings of their transcripts, with comparable results. Thus, the FIDO descriptors themselves seemed adequate for capturing both the target

TABLE 4
Summary of FIDO Data by Six Scoring Methods ($N = 167$)

Scoring method	Effect size	Composite z-score	Probability	# Trials $p < .05$*	% Trials $p < .05$*	% Trials $p < .50$
FIDO	0.1343	1.735	.041	10 (8)	6% (5%)	54%
Binary	0.0761	0.984	.163	13 (12)	8% (7%)	53%
Ternary	0.1598	2.065	.019	5 (6)	3% (4%)	56%
Exact	0.1495	1.932	.027	17 (6)	10% (4%)	54%
Distributive	0.1453	1.878	.030	12 (6)	7% (4%)	57%
Weighted distributive	0.1467	1.896	.029	15 (6)	9% (4%)	55%

* Numbers in parentheses indicate number of trials with negative z-scores, $p < .05$.

information and the percipients' imagery. The diminishment of the yield evidently had its source elsewhere.

VIII. Distributive Scoring

Shortly after completion of the FIDO analyses, an REG-based human/ machine study had indicated that operator pairs of opposite sex, working together with a shared intention, produced substantially stronger effects than same-sex pairs or individual operators.[35] This, in turn, had led to a comprehensive examination of nine of PEAR's human/machine databases, which were found to display significant gender-related differences in individual operator achievement.[36] Although hints of possible gender-related trends had also been noted in the PRP data, the previous pool of contributing percipients and agents had been too small and disproportionately balanced to determine whether such gender-pairing might be a significant factor in these experiments as well. To explore this hypothesis, a new body of remote perception experiments was performed using a balanced pool of same- and opposite-sex participant pairs, each contributing an equal number of trials.

This new protocol required each percipient/agent pair to generate a series consisting of five trials. Ideally, the same pair would produce another five-trial series with their roles reversed. Since a concern had been raised that providing feedback to participants at the conclusion of each trial could introduce a possible bias in subsequent trials, feedback to participants was withheld until all five trials of a series were completed, and each target selected from the pool in instructed experiments was returned before the next trial. To preclude any possibility of shared response bias, all analyses were based solely on local subset comparisons within a given series.

As an added attempt to improve the scoring methodology, a new descriptor check sheet was designed that permitted participants to respond to each question on a distributive scale of 0 to 9 to indicate the relative prominence of each of 30 descriptor features. Similar to the prior methods, the results were evaluated by

constructing a 5×5 matrix for each series by scoring every target against every perception. These individual scores, in turn, were drawn from various 10×10 matrices that cross-indexed and assigned values to every possible pair of 0–9 descriptor rankings. Again, several different recipes were applied:

- A direct-match matrix that awarded a score of 1 for any exact descriptor match and 0 for any mismatch.
- A binary matrix that treated any response of 0–4 as a "no," and any response of 5–9 as a "yes," with a correct match assigned a score of 1 and an incorrect match a score of 0.
- A ternary matrix that treated 0–2 as a "no," 3–6 as an "unsure," and 7–9 as a "yes," and assigned a score of 2 to any correct "yes" or "no" match, 1 to a correct "unsure" match, and 0 to any other response.
- A distributive matrix that assigned a score of 2 for a direct match, 1 for a mismatch by one or two levels in the descriptor rankings, and 0 for any other mismatches.
- An extended distributive matrix that assigned a score of 10 to a direct match, 5 to an adjacent match, 2 to a response two points removed from the correct rank, 1 to a response three points removed, and 0 to any other response.
- A weighted distributive matrix that assigned scores of 9 for direct matches at the extremes of the range (0 or 9), with decreasing credit as the match approached the middle of the range; *i.e.*, correct matches of 1 or 8 received a score of 8, matches of 2 or 7 received a 7, etc. Scoring for adjacent matches followed a similar pattern of reduced credit as the rank approached the middle of the range.

As before, the sum of the individual descriptor scores constituted the total score for a given trial, and the scores of the five matched trials were compared with those of the 20 mismatched scores to determine the statistical merit of each series.

Thirty experimental series comprising 150 trials were generated using this distributive protocol by 12 participant pairs, 8 of whom produced at least two series together with the percipient/agent roles reversed. The results are summarized in Table 5.

Once again, there was reasonably good agreement among the six scoring recipes, but the overall results were now completely indistinguishable from chance. No more than the expected number of significant trials emerged in the analyses, and the low statistical resolution in defining the local empirical chance backgrounds, a consequence of the small size of the scoring matrices, made calculation of individual trial z-scores virtually meaningless. In a certain sense, this was reminiscent of one of the problems that had stimulated development of the analytical judging methodologies 18 years earlier, namely, the statistical inefficiency of assessing the informational content of individual trials in small experimental series. But now the phenomenon itself seemed to

TABLE 5
Summary of Distributive Data by Six Scoring Methods (30 Series, 150 Trials)

Scoring method	Effect size	Composite z-score	Probability	# Series p < .05*	# Trials p < .05*	% Trials p < .05*	% Trials p < .50
Direct match	−0.0088	−0.108	.543	2 (0)	6 (6)	4% (4%)	46%
Binary	−0.0684	−0.838	.799	0 (1)	8 (3)	5% (2%)	47%
Ternary	−0.0342	−0.419	.662	0 (0)	5 (5)	3% (3%)	55%
Distributive	−0.0501	−0.613	.730	1 (0)	5 (5)	3% (3%)	51%
Extended distributive	−0.0745	−0.912	.819	1 (0)	6 (9)	4% (6%)	52%
Weighted distributive	−0.0394	−0.483	.685	2 (0)	6 (8)	4% (5%)	53%

* Numbers in parentheses indicate number of trials with negative z-scores, $p < .05$.

have disappeared. And given the lack of any statistical yield in these data, it was not possible to ascertain whether there was any evidence of co-operator or gender differences, the question that had originally prompted this exploration.

In pondering this paradox, we became cognizant of a number of subtler, less quantifiable factors that also might have had an inhibitory effect on the experiments, such as the laboratory ambience in which the experiments were being conducted. For example, during the period in which the FIDO data were being generated, we were distracted by the need to invest a major effort in preparing a systematic refutation to an article critical of PEAR's earlier PRP program.[37,38] Although most of the issues raised in that article were irrelevant, incorrect, or already had been dealt with comprehensively elsewhere and shown to be inadequate to account for the observed effects,[23] this enterprise deflected a disproportionate amount of attention from, and dampened the enthusiasm for, the experiments being carried out during that time. Beyond this, in order to forestall further such specious challenges, it led to the imposition of additional unnecessary constraints in the design of the subsequent distributive protocol. Although it is not possible to quantify the influence of such intangible factors, in the study of consciousness-related anomalies where unknown psychological factors appear to be at the heart of the phenomena under study, they cannot be dismissed casually.

IX. Review and Discussion

The evidence acquired in the early remote perception trials had raised profound questions in the minds of the PEAR researchers, similar, no doubt, to those of the countless others who, over the course of history, had experienced first-hand the validity of Paracelsus' remarkable claim. The possibility that ordinary individuals can acquire information about distant events by these inexplicable means, even before they take place, challenges some of the most fundamental premises of the prevailing scientific worldview. PEAR's efforts to devise strategies capable of representing the information acquired in the

remote perception process in a manner amenable to quantitative analysis had followed the traditional scientific method, *i.e.*, to design experiments capable of reproducing the phenomenon under carefully controlled conditions, to systematically eliminate sources of extraneous noise in order to bring the phenomenon in question into sharper focus, and to pose theoretical models to dialogue with these empirical results.

The early phases of the program provided encouraging indications that this could be accomplished via a set of standardized descriptor queries, addressed to both the agent's description of the physical target and to the percipient's stream-of-consciousness narrative, that would serve as an ''information net'' to capture the essence of the anomalous communication. *Ex post facto* application of this technique to existing data seemed to confirm the efficacy of this approach, producing results that were consistent with previous human judge assessments and encouraging continued explorations. In the second phase of the program, *ab initio* utilization of this method in a new body of experiments also produced highly significant results. While the average effect size of these was somewhat smaller than that of the original *ex post facto* subset, this was attributed primarily to the fact that these were the data on which the descriptor questions and analytical techniques had been based. Nevertheless, the statistical yield of the *ab initio* data still was sufficiently robust to indicate that the new method could serve its intended purpose adequately.

Yet, like so much of the research in consciousness-related anomalies, replication, enhancement, and interpretation of these results proved elusive. As the program advanced and the analytical techniques became more sophisticated, the empirical results became weaker. It appeared as if each subsequent refinement of the analytical process, intended to improve the quality and reliability of the ''information net,'' had resulted in a reduction of the amount of raw information being captured. This diminution of the experimental yield prompted extensive examination of numerous factors that could have contributed to it. After exploring and precluding various possible sources of statistical or procedural artifact, however, we were forced to conclude that the cause of the problem most likely lay somewhere in the subjective sphere of the experience.

Throughout the course of the program, when participants had been queried about their personal reactions to the encoding process, their most common complaint was a feeling of being ''constrained'' by the required forced-choice binary queries. In response, the FIDO phase was implemented to permit participants more freedom in formulating their responses. Although the FIDO database appeared to contain a considerable number of impressionistically successful trials, the composite quantitative results now were only marginally significant.

The failure of FIDO to reinvigorate the PRP program, plus the desire to examine variations in individual performance, led to yet another encoding strategy with even more response flexibility, *i.e.*, the distributive methodology. Although this method was intended to alleviate participants' feelings of subjective

constraint, concerns about the possibility of participant response biases imposed additional procedural restrictions. It was evident from the null results of the 150 distributive trials that all efforts to enhance the effect by progressively more elaborate analysis techniques not only had failed, but even had proven counter-productive. Although the judging methodology had been proven to serve its intended analytical purpose, the progressive attenuation of the yield suggested that there was some kind of interference taking place between the analytical measures and the generation of the effects they were attempting to measure.

The trend is clearly evident on re-examination of the cumulative deviation graph of Figure 1, which plots chronologically the cumulative results of all 336 formal binary-encoded trials and displays a potentially instructive clue to the inexorable decrease in effect size. Following the initial sharp slope representing the strong yield of the original 59 *ex post facto* trials, the slope of the subsequent 277 *ab initio* trials can be seen to consist of two distinct segments. The first of these, comprising the initial 168 *ab initio* trials (60 through 227 on the *x*-axis) has a consistent positive slope, albeit shallower than that of the earlier *ex post facto* data. The slope of the second segment (trials 228 through 336), which consists of the 109 trials from the second phase of the *ab initio* experiments, is noticeably flatter. The beginning of this second segment would therefore appear to be the point at which the experimental yield began to deteriorate. Figure 4 plots the comparative effect sizes of the data from these various experimental periods, reconfirming the systematic decrease of the yield beginning with the second phase of the *ab initio* binary experiments. The numerical results of these segments are presented in Table 6. (Again, the effect sizes displayed in the graph and table were calculated by dividing the *z*-scores for each database by the square root of the number of trials in that subset, and thus indicate the average *z*-score per trial.)

While the composite yield of the total database remains highly significant, it is evident that this result is driven primarily by the much stronger yields of the earlier trials, bolstered by the substantial size of the overall database itself. The success of the analytical judging technique in the early phases of the program, and its apparent insensitivity to the particular scoring matrices invoked, confirms that such an approach can indeed be deployed successfully as a strategy for quantifying this inherently subjective process. Nonetheless, something clearly changed in the second phase of the *ab initio* experiments that resulted in a substantial weakening of the effect being quantified. Since both phases of the *ab initio* portion of the program utilized identical descriptor questions and scoring algorithms, their analytical effectiveness therefore can be ruled out as the source of the lower yield in the later phases of the program.

Another pattern became evident when we returned to the raw free-response data with this in mind. The free-response descriptions in the later trials were considerably shorter than those generated in the earlier ones, some of which had run to several pages of narrated perceptions. Indeed, in many of these later trials, percipients' verbal descriptions consisted of only a few cursory phrases, intended
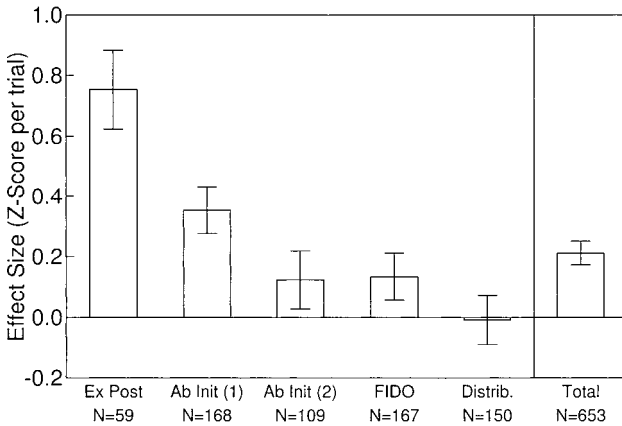
Fig. 4.   Effect sizes of various data subsets.

TABLE 6
PRP Summaries by Database

Database	# Trials	# Series	# Participants*			Composite z-score	Effect size	Probability
			# Agents	# Percipients	Total			
Ex post facto	59	7	4	13	16	5.792	.754	3×10^{-9}
Ab initio	277	42	13	26	30	4.378	.263	6×10^{-6}
Initial trials	168	29	9	21	23	4.582	.354	2×10^{-6}
Later trials	109	13	7	13	15	1.291	.124	.098
FIDO	167	9	19	22	25	1.735	.134	.041
Distributive	150	30	15	15	16	−0.108	−.009	.543
TOTAL	653	88	39	59	69	5.418	.212	3×10^{-8}

* Some individuals contributed to more than one database, in both percipient and agent capacity.

simply to clarify nuances of their descriptor responses, and provided little in the way of the stream-of-consciousness imagery they had been asked to generate. It appeared that as the percipients became more familiar with the descriptor questions, their subjective impressions were increasingly guided and circumscribed by them, as though the questions were establishing the informational framework for their responses. The original free-response remote perception experiment thus had taken on the characteristics of a multiple-choice task, and the locus of the experience had shifted from the realm of intuition to that of intellect.

X. From Analysis to Analogy

Having exhausted the search for the source of the remote perception signal deterioration in the analytical techniques themselves, we are driven to look

further afield for a satisfactory explanation. If we step back to review the program from a broader perspective, we note that all of the methodological ''improvements'' introduced to refine the scoring techniques had been directed toward more efficient extraction of the anomalous information and elimination of possible sources of artifact or bias. Some were efforts to achieve ''sharper definition'' of the remote perception ''signal,'' others were attempts to ''tighten'' the experimental ''controls,'' and a few were designed to ''clarify'' certain characteristics of the communication ''channel.'' All these terms reflect an emphasis on achieving increasingly precise specification and reducing the noise or uncertainty in the process. Yet, each increment of analytical refinement appears to have resulted in a systematic reduction not of the ''noise'' but of the ''signal'' itself. This raises the somewhat radical possibility that manifestation of the anomaly may actually require a certain degree of the very noise, or uncertainty, that we had invested so much effort to reduce. It is a possibility, however, for which precedent can be found in other domains of scholarly inquiry, and is therefore worth consideration in the present context.

The most immediate technical examples of this complementarity of signal and noise are the human/machine experiments carried out in our laboratory and elsewhere.[39] All of these studies employ some form of random processor, and the anomalous effects appear as departures of their random outputs from chance expectation. It is as if the ''noise'' of the random process provides the essential raw material out of which the mind of the operator is able to construct a small amount of ordered ''signal.''

Such effects are by no means restricted to explicit anomalies research. Similar departures from canonical expectations can be found in contemporary engineering applications of ''stochastic resonance,'' wherein a deliberate increase in the overall level of noise in certain kinds of lasers or sensitive electronic circuits can actually enhance the detection of weak, fluctuating signals.[40,41] Other studies have demonstrated that the introduction of an element of chaos into certain types of nonlinear processes, such as the interaction of two otherwise independent random oscillators, can stimulate synchronous behavior between the transmitter and the receiver.[42,43] In each of these instances, information or order has been introduced into a sensitive nonlinear physical system, not by reducing the ambient noise, but by *increasing* it.

Of particular interest for our purpose is the researchers' unanticipated observation that in such synchronization processes the receiver actually recorded changes in the signal *before* the transmitter recorded the transmission of those changes. In other words, the system seemed capable of anticipating the synchronization. The engineers who carried out the studies remarked that, ''We would thus expect that any of those analogous systems which exhibit chaos should also be liable to anticipating synchronization. We thus hope that our work will act as a stimulus to explore the opportunities for observing anticipating synchronization in physical, chemical, biological and socio-economic systems.''[41] Following this suggestion, we might note that, in

a certain sense, the remote perception process qualifies as an example of a "sensitive nonlinear system with a weak fluctuating signal" that exhibits a certain degree of chaos, and that the participants in these experiments function as "two otherwise independent random oscillators." Hence, it well may be that our signal is also dependent upon a background of random noise for its manifestation. If so, it would appear that it was our attempts to enhance the remote perception signal by sharpening the specificity of the information channel that could, in fact, have been responsible for the attenuation of the signal.

Reaching farther afield for relevant analogies, the accepted model of biological evolution incorporates the importance of uncertainty in enhancing information. Darwinian theory postulates that living species adapt to their environment by selecting for specific traits that emerge in the process of random genetic mutation. This process is itself strongly dependent on the generation of "noise" emerging from the massive redundancy of continuously recombined genetic information. When the randomness of this process is limited, as in repeated interbreeding, the short-term advantage of increased predictability of inherited traits is offset by longer-term weakening of the genetic strain of the species.

Insights can also be derived from a quite different realm of human experience, namely, the practice of certain mystical divinatory traditions where anomalous relationships between signal and noise are also evident. In most of these, a clearly defined question is submitted to some kind of random process for the purpose of accessing information unavailable to the conscious mind. Typically, the response comes in imprecise or symbolic form that requires translation into meaningful or pertinent terms. One such example is the renowned Oracle of Apollo at Delphi in ancient Greece, a highly respected source of wisdom that long played a central role in Greek culture and politics. Consultation of the oracle involved a priestess called the Pythia who, crowned in laurel and in an altered state of consciousness stimulated by vapors arising from a cleft in the earth over which she sat on a tripod, produced a "free response" utterance, which was then interpreted by the attending priest in response to the seeker's query. Two points of potential relevance here are the non-analytical, receptive state of mind of the "percipient," and the deferment of interpretation by the "judge" until *after* the experience has been completed.

Another ancient oracle, still widely used, is the Chinese "Book of Changes," or *I Ching*, a divination process that involves generation of a sequence of random binary events, the results of which are represented as two "trigrams." These are referred to a table, or matrix, that identifies each of the 64 possible combinations, or "hexagrams," with a specific text that is then consulted to obtain a response to the original query. Notwithstanding the subjective nature of the interpretation of the texts, a vast body of evidence accumulated over many millennia testifies to the efficacy of the *I Ching* in producing accurate and consequential results. Despite the claim of many rationalists that such oracles are nothing more than bizarre combinations of wishful thinking and "mere

chance,'' this is the same ''irrational'' formula that seems to underlie the remote perception phenomena that have now been demonstrated, by rigorous analytical quantification, to convey more meaningful information than can be attributed to ''mere chance.'' Hence the principles invoked by the ancient sages in developing the *I Ching* may shed some light on these more contemporary anomalies.

Psychologist Carl Jung, who devoted more than 30 years to the study of the *I Ching*, pointed out in his Foreword to the classic Richard Wilhelm translation[44] that ''we know now that what we term natural laws are merely statistical truths and thus must necessarily allow for exceptions. . . . If we leave things to nature, we see a very different picture: every process is partially or totally interfered with by chance, so much so that under natural circumstances a course of events absolutely conforming to specific laws is almost an exception.'' He relates the emphasis placed by the ancient Chinese mind on chance and the subjective interpretation of events to the modern world of quantum mechanics, where the reality of inherently random microscopic physical events includes the observer as well as the observed. In both domains, what Jung refers to as the ''hidden individual quality of things and men'' draws on the unconscious and intangible qualities that undergird the experiences of the conscious mind and the tangible physical world, respectively, in similar fashion to the conceptual framework described in our paper, ''A Modular Model of Mind/Matter Manifestation (M^5).''[45] Both Jung's representation and our own emphasize that the causal and synchronistic perspectives of reality are complementary, rather than mutually exclusive. Jung maintains that the ''coincidence'' of a synchronistic event occurs ''because the physical events are of the same quality as the psychic events and because all are the exponents of one and the same momentary situation.''[44] Our representation of this concept speaks of the emergence of both cognitive experience and physical events from a common underlying substrate of the unconscious mind and the undifferentiated world of physical potentiality, wherein the distinction between mind and matter blurs into uncertainty. Given their common origin, it should not be surprising to observe correlations between their manifested expressions in the worlds of mental and physical ''reality.'' Just as the concept of complementarity in quantum mechanics brings with it a certain degree of uncertainty that makes it impossible to achieve absolute precision in two frames of reference simultaneously, the complementarity of an ''objective'' causal picture of reality and a ''subjective'' synchronistic one also may necessitate tolerance of a degree of uncertainty in both dimensions.

In many respects, the empirical evidence from remote perception, as well as from other domains of anomalies research, is more compatible with an acausal, or synchronistic, model than with a causal one. Although we have recognized this in principle, our experimental approach and the language we have deployed in describing the effects has betrayed certain causal assumptions. For example, despite repeated comments from participants that the PRP experience felt more like ''sharing'' than ''sending and receiving,'' we persisted in speaking of information ''transmission.'' Similarly, our enduring efforts to extract the

"signal" from the "noise" also reflected a more deterministic orientation. Yet, Jung's model, the ancient divinatory traditions, evolutionary theory, contemporary signal processing research, and human/machine anomalies all suggest that noise may be a requisite component of the process of signal generation, and that objective linear causality may not prevail under these circumstances.

If one defines "noise" in the remote perception context as the percipient's uncertainty, or lack of conscious knowledge, about the target, and "signal" as the content of valid information acquired in the process, these diverse analogies can be quite instructive. For example, the early experiments, wherein percipients were asked simply to generate an unfocused, free-response stream of consciousness, were in this sense more "noisy" than the later efforts, where percipients' imagery was guided by a more structured information "grid" or "filter" of descriptor queries. In those trials that were only encoded *ex post facto*, the participants had no knowledge of the information filter that would be imposed only well after the data were generated, and they seemed more easily able to access information about the targets. In the first generation of *ab initio* binary-encoded trials, when descriptor check-sheets were something of a novelty and percipients were still urged to generate their free-response descriptions before attempting descriptor encoding, the transcripts tended to be somewhat shorter, but most of them still comprised a free-association type of narrative. These trials also produced highly successful results, albeit of a somewhat smaller average effect size. By the time of the later *ab initio* experiments, however, when we had acquired greater confidence in the efficacy of the analytical judging approach, less importance was placed on the raw free-response data and this shift of emphasis was reflected in the abbreviated, even cursory, percipient responses. In retrospect, it is apparent from the content of these shorter transcripts that the percipients were anticipating the descriptor questions and inadvertently focusing their attention on those particular aspects of their experience. Although the intent of the quaternary, and then distributive, descriptor questions was to relieve the participants' sense of "constraint," these more complex forms of questions appear to have had the opposite effect, forcing percipients to pay even more attention to the nuances of the information grid and thus filtering out any signal that was not perceived to be "relevant." In this way, the background "noise" was reduced even further, and more structured cognitive processes, associated with achieving internal consistency in what had essentially become a forced-choice task, effectively restricted the flow of unconscious imagery.

It is also telling that, until recently, this trend had not even been perceived as a problem by the researchers. Typing 30 numbers into a computer was much easier than the task of evaluating lengthy verbal transcripts, and the ability to acquire a quantitative indication of the merit of an individual trial increasingly replaced the spontaneous excitement of finding apparent correspondences in the raw data. The shift in experimental perspective from predominantly subjective to almost totally analytical was so gradual that little consideration was given to

the possible costs of such a transition. For example, combination of the data from the first and second phases of the *ab initio* experiments was justified solely on technical grounds, with no serious consideration given to the implications of a change from ranking the quality of a trial to measuring its specific information content, other than the relative efficiency and statistical power of the two approaches. The subsequent effort expended on refining the technical and analytical components of the program, rather than on trying to understand what the participants were really trying to tell us when they complained of feeling "constrained" by the descriptor questions, further exacerbated the overemphasis on quantitative precision that ultimately may have suffocated the subtle, but essential, subjective signal.

The larger effect size of the "instructed" *vs.* the "volitional" trials also supports the importance of retaining an adequate component of noise or uncertainty in the system. When percipients attempted to describe scenes chosen by a random process that precluded utilization of any prior knowledge about the agent's habits or personal preferences, their perceptions contained a larger component of anomalous information. In the volitional protocol, where one might imagine a certain *a priori* advantage, percipients' rational expectations may have imposed yet another kind of information filter that inhibited the subtle "signal detection" process. In other words, the strongest "signals" appear to have been generated under the "noisiest" conditions, *i.e.*, in the absence or minimization of any orderly or rational form of structural information. (It may be interesting to note in this regard that approximately 66% of the *ab initio* binary trials, 98% of the FIDO trials, and 77% of the distributive trials followed the volitional protocol, whereas 53% of the *ex post facto* trials were instructed.)

One might even speculate that the overall success of these experiments derives in considerable measure from the "irrational" nature of the remote perception task itself. When requested to describe a spatially and temporally remote scene without access to any known sensory channel, percipients are forced to abandon any rational strategy for fulfilling such an assignment. With cognitive functioning thus confounded by uncertainty, leaving the conscious mind less able to mask the subtle signal with rational associations, the unconscious mind of the percipient may better be able to access the "hidden individual quality of things and men."

Although a degree of uncertainty may indeed be necessary for the generation of remote perception effects, the complementary relationship between signal and noise we are proposing nevertheless requires retention of a comparable dimension of structure in the process. Recall, for example, that the early exploratory trials, where percipients did not know the identity of the agent or the time of target visitation, produced completely null results (Table 3). As in the *I Ching* or other divinatory arts, where it is essential that the querant pose a clearly defined question, the remote perception process also seems to require the percipient to establish some minimal "boundary conditions" when addressing the unknown target. If indeed such a process involves an excursion into the

unconscious realm of undifferentiated potential in order to acquire specific information, some corresponding specific question would appear to be a prerequisite. To complement this facilitative function, some form of quantitative assessment of the amount of anomalous information is indispensable if the study of remote perception is to qualify as a scientific enterprise.

To this end, we have proposed in several previous publications that a more astute balance between the analytical and the aesthetic dimensions of such phenomena needs to guide any future explorations of consciousness-related anomalies.[32,45–49] In the article entitled ''Science of the Subjective,''[49] we observed how ''in the interplay of objective intellect and subjective spirit, we are dealing with the primordial conjugate perspectives whereby consciousness triangulates its experience.'' This complementary relationship has now been confirmed in the record of our remote perception research. That is, the subjective spirit of these experiences appear to be more effectively attained when unencumbered by analytical or cognitive overlays and its inherent uncertainties are both acknowledged and utilized. However, the equally important role of objective intellect must serve to enhance, rather than to inhibit, the process and our eventual understanding of it.

Appendix A

Local Descriptor Probabilities and Individual Performance

The scores presented in the summaries of Table 2 had been calculated using the local *a priori* probabilities associated with each subset, following the same procedure that had been deployed for all of the major analyses in the first phase

of the analytical judging program.[25(Appx.C)] Those early explorations had established that when the local *a priori* probabilities were used to score a particular subset using a given scoring method, the empirical chance distributions resulting for different subsets appeared to be statistically indistinguishable. It thus had been concluded that a single empirical chance distribution, namely the one resulting from the largest assembly of formal data, could be used as a reliable reference standard for any subset, provided that the subset's trial scores were computed using its own local *a priori* probabilities.

Unfortunately, this uniformity of chance distributions is only approximately correct. A re-evaluation of this technique illustrated a mechanism whereby internal variations in the *a priori* probabilities among different subsets of the database could potentially produce artificially inflated, or deflated, scores in the matched-trial distributions relative to the off-diagonal population of mismatches. For example, a given percipient/agent pair might happen to share a similar encoding style, such as a tendency to respond affirmatively to ambiguous features, or particular preferences for certain descriptors, which could result in their trials having responses that were more closely correlated than those of the mismatched scores constituting the reference distribution. Similar biases also might arise from geographical or seasonal variations, or other possible causes.

Since the apparent indistinguishability of the chance distribution for a number of large data subsets cannot be guaranteed theoretically, it is necessary to verify empirically that the overall results are not in fact spuriously inflated by such biasing mechanisms. The possible influence of idiosyncratic individual patterns of *a priori* response probabilities in agent and percipient encoding styles was examined using the data produced by the 29 agent/percipient pairs who had contributed five or more trials to the composite database. (Collectively, these 29 pairs were responsible for 274 of the 336 formal trials.) The results of this test for local biasing are shown in Figure A, which displays an array of traces for these 274 trials, after the style of Figure 1. The individual plotted points are the cumulative *z*-scores achieved by each of the 29 agent/percipient pairs based on three distinct calculation methods. The "non-local" method calculates each trial score using the *a priori* probabilities for the full formal database and computes its *z*-score against the standard empirical chance distribution for the overall database. In other words, this trace is simply the composite *z*-score assigned to the subset of trials contributed by given agent/percipient pairs, extracted from the results of the overall database of 336 formal trials. In comparison, the "local alpha" score is derived by scoring each percipient/agent pair's contributions on the basis of its own internal *a priori* probabilities, but still referring these scores to the overall empirical chance distribution. The "local distribution" calculation removes all reference to global distributions, and along with it any possibility of local-biasing effects, by scoring each agent/percipient pair's data not only with its own local *a priori* probabilities, but against its own local mismatch distribution.
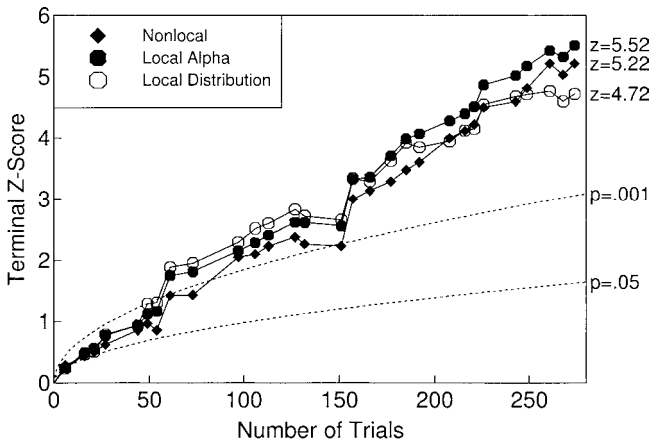
Fig. A.   Cumulative *z*-score progress for three alternative scoring techniques.

With few exceptions, all of which are associated with very small datasets, the three scoring strategies produce a reassuring degree of agreement, especially in the composite yields. It is evident from Figure A that these three methods are not statistically distinguishable, and that any inflation or deflation of the overall effect due to local biasing is less than the inherent statistical uncertainty of the scoring procedure. It therefore may be concluded that, within the limits of the statistical resolution, encoding artifact is not a significant contributor to these experimental results.

The rank-ordered effect sizes obtained by each of the 28 percipients and 15 agents who contributed more than one trial to the database were also examined. Some 25% of the percipients, 40% of the agents, and 21% of the percipient/agent pairs produced statistically significant overall results, whereas only 5% of each group would be expected to do so by chance. All but two percipients and two agents generated net positive effects, compared to the 50% chance expectation, and of these four individuals, three produced positive results when functioning in the alternate role. A separate data subset, consisting of only the first trials from each of the 38 percipients contributing to the formal database, was also calculated to examine the possibility that the composite yield might have been distorted by large databases produced by any given percipient. Despite the small size of this group of trials, the results display the same linear consistency as the full database, achieving a highly significant composite *z*-score of 3.890. Thus, it is also clear that the success of the overall results is not attributable to exceptional performance by only a few participants.[25]

References

1. Hartmann, F. (1973). *Paracelsus: Life and Prophecies.* Blauvelt, NY: Rudolf Steiner (p. 105).
2. Targ, R., & Puthoff, H. E. (1974). Information transmission under conditions of sensory shielding. *Nature, 252,* 602–607.

3. Puthoff, H. E., & Targ, R. (1976). A perceptual channel for information transfer over kilometer distances: Historical perspective and recent research. *Proceedings of the IEEE*, *64*, 329–354.
4. Allen, S., Green, P., Rucker, K., Cohen, R., Goolsby, C., & Morris, R. L. (1975). A remote viewing study using modified version of the SRI procedure. In Morris, J. D., Roll, W. G., & Morris, R. L. (Eds.), *Research in Parapsychology*. Metuchen, NJ: Scarecrow Press (pp. 46–48).
5. Hastings, A. C., & Hurt, D. B. (1976). A confirmatory remote viewing experiment in a group setting. *Proceedings of the IEEE*, *64*, 1544–1545.
6. Rauscher, E. A., Weissmann, G., Sarfatti, J., & Sirag, S. P. (1976). Remote perception of natural scenes, shielding against ordinary perception. In Morris, J. D., Roll, W. G., & Morris, R. L. (Eds.), *Research in Parapsychology*. Metuchen, NJ: Scarecrow Press (pp. 41–45).
7. Vallee, J., Hastings, A. C., & Askevold, G. (1976). Remote viewing experiments through computer conferencing. *Proceedings of the IEEE*, *64*, 1551–1552.
8. Whitson, T. W., Bogart, D. N., Palmer, J., & Tart, C. T. (1976). Preliminary experiments in group ''remote viewing.'' *Proceedings of the IEEE*, *64*, 1550–1551.
9. Puthoff, H. E., & Targ, R. (1977). *Mind Reach*. New York: Delacorte.
10. Chotas, H. (1978). Remote viewing in the Durham area. *Journal of Parapsychology*, *42*, 61–62.
11. Tart, C. T., Puthoff, H. E., & Targ, R. (Eds.). (1979*). Mind at Large: IEEE Symposia on the Nature of Extrasensory Perception*. New York: Praeger Special Studies.
12. Karnes, E. W., Sussman, E. P., Klusman, P., & Turcotte, L. (1980). Failures to replicate remote viewing using psychic subjects. *Zetetic Scholar*, *6*, 66–76.
13. Schlitz, M., & Gruber, E. (1980). Transcontinental remote viewing. *Journal of Parapsychology*, *44*, 305–317.
14. Puthoff, H. E., Targ, R., & May, E. C. (1981). Experimental psi research: Implications for physics. In Jahn, R. G. (Ed.), *The Role of Consciousness in the Physical World*. Boulder, CO: Westview Press (pp. 37–86).
15. May, E. C., Utts, J. M., Humphrey, B. S., Luke, W. L. W., Frivoid, T. J., & Trask, V. V. (1990). Advances in remote-viewing analysis. *Journal of Parapsychology*, *54*, 193–228.
16. Nelson, R. D., Dunne, B. J., Dobyns, Y. H., & Jahn, R. G. (1996). Precognitive remote perception: Replication of remote viewing. *Journal of Scientific Exploration*, *10*, 109–110.
17. May, E. C. (1996). The American Institutes for Research review of the Department of Defense's STAR GATE program: A commentary. *Journal of Scientific Exploration*, *10*, 89–108.
18. Puthoff, H. E. (1996). CIA-initiated remote viewing program at Stanford Research Institute. *Journal of Scientific Exploration*, *10*, 63–76.
19. Targ, R. (1996). Remote viewing at Stanford Research Institute in the 1970s: A memoir. *Journal of Scientific Exploration*, *10*, 77–88.
20. Bremseth, Commander L. R. (2001). *Unconventional Human Intelligence Support: Transcendent and Asymmetric Warfare Implications of Remote Viewing*. Graduate thesis, Marine Corps War College, Marine Corps University, Marine Corps Combat Development Command, Quantico, VA.
21. Puthoff, H. E. (2001). CIA-initiated remote viewing at Stanford Research Institute. *The Intelligencer: Journal of U.S. Intelligence Studies*, *12*, 60–67.
22. Jahn, R. G., Dunne, B. J., & Jahn, E. G. (1980). Analytical judging procedure for remote perception experiments. *Journal of Parapsychology*, *44*, 207–231.
23. Jahn, R. G., Dunne, B. J., Nelson, R. D., Jahn, E. G., Curtis, T. A., & Cook, I. A. (1982). Analytical judging procedure for remote perception experiments—II: Ternary coding and generalized descriptors. Technical Note PEAR 82002, Princeton Engineering Anomalies Research, School of Engineering/Applied Science, Princeton University, Princeton, NJ.
24. Dunne, B. J., Jahn, R. G., & Nelson, R. D. (1983). Precognitive remote perception. Technical Note PEAR 83003, Princeton Engineering Anomalies Research, School of Engineering/Applied Science, Princeton University, Princeton, NJ.
25. Dunne, B. J., Dobyns, Y. H., & Intner, S. M. (1989). Precognitive remote perception III: Complete binary data base with analytical refinements. Technical Note PEAR 89002, Princeton Engineering Anomalies Research, School of Engineering/Applied Science, Princeton University, Princeton, NJ.
26. Dunne, B. J., & Jahn, R. G. (2002). Information and uncertainty: 25 years of remote perception research. Technical Note PEAR 2002.01, Princeton Engineering Anomalies Research, School of Engineering/Applied Science, Princeton University, Princeton, NJ.
27. Dunne, B. J., & Bisaha, J. P. (1979). Precognitive remote viewing in the Chicago area: A replication of the Stanford experiment. *Journal of Parapsychology*, *43*, 17–30.

28. Bisaha, J. P., & Dunne, B. J. (1979). Multiple subject and long-distance precognitive remote viewing of geographical locations. In Tart, C. T., Puthoff, H. E., & Targ, R. (Eds.), *Mind at Large: IEEE Symposia on the Nature of Extrasensory Perception*. New York: Praeger Special Studies (pp. 109–124).

29. Solfvin, G., Kelly, E., & Burdick, D. (1978). Some new methods for analysis for preferential ranking data. *Journal of the American Society for Psychical Research*, *72*, 93–109.

30. Dunne, B. J. (1979). *Precognitive Remote Perception: A Critical Overview of the Experimental Program*. Master's thesis, Committee on Human Development, University of Chicago, Chicago, IL.

31. Andresen, J. (1984). *Statistical Tests of Experimental Remote Perception Trials*. Undergraduate thesis, Department of Civil Engineering, School of Engineering/Applied Science, Princeton University, Princeton, NJ.

32. Jahn, R. G., & Dunne, B. J. (1987). *Margins of Reality: The Role of Consciousness in the Physical World* (Section III). San Diego, CA: Harcourt Brace Jovanovich.

33. Jahn, R. G., Dunne, B. J., & Nelson, R. D. (1987). Engineering anomalies research. *Journal of Scientific Exploration*, *1*, 21–50.

34. Dunne, B. J., & Jahn, R. G. (1992). Experiments in remote human/machine interaction. *Journal of Scientific Exploration*, *6*, 311–332.

35. Dunne, B. J. (1993). Co-operator experiments with an REG device. In Rao, K. R. (Ed.), *Cultivating Consciousness for Enhancing Human Potential, Wellness, and Healing*. Westport, CT: Praeger (pp. 111–121). (Also available as Technical Note PEAR 91005, Princeton Engineering Anomalies Research, School of Engineering/Applied Science, Princeton University, Princeton, NJ.)

36. Dunne, B. J. (1998). Gender differences in human/machine anomalies. *Journal of Scientific Exploration*, *12*, 3–55.

37. Hansen, G. P., Utts, J., & Markwick, B. (1992). Critique of the PEAR remote-viewing experiments. *Journal of Parapsychology*, *56*, 97–113.

38. Dobyns, Y. H., Dunne, B. J., Jahn, R. G., & Nelson, R. D. (1992). Response to Hansen, Utts, and Markwick: Statistical and methodological problems of the PEAR remote viewing [*sic*] experiments. *Journal of Parapsychology*, *56*, 115–146.

39. Jahn, R. G., Dunne, B. J., Nelson, R. D., Dobyns, Y. H., & Bradish, G. J. (1997). Correlations of random binary sequences with pre-stated operator intention: A review of a 12-year program. *Journal of Scientific Exploration*, *11*, 345–367.

40. McNamara, B., Wiesenfeld, K., & Roy, R. (1988). Observation of stochastic resonance in a ring laser. *Physical Review Letters*, *60*, 2626–2629.

41. McNamara, B., & Wiesenfeld, K. (1989). Theory of stochastic resonance. *Physical Review A*, *39*, 4854–4869.

42. Jones, R. J., Rees, P., Spencer, P. S., & Shore, K. A. (2001). Chaos and synchronisation of self-pulsating laser diodes. *Journal of the Optical Society of America*, *B*, *18*, 166–172.

43. Sivaprakasam, S., Shahverdiev, E. M., Spencer, P. S., & Shore, K. A. (2001). Experimental demonstration of anticipating synchronization in chaotic semiconductor lasers with optical feedback. *Physical Review Letters*, *87*, 154101.

44. Jung, C. G. (1950). Foreword to *The I Ching* (R. Wilhelm, trans.). Princeton, NJ: Princeton University Press, pp. *xxxi–xxxix*.

45. Jahn, R. G., & Dunne, B. J. (2001). A modular model of mind/matter manifestation (M^5). *Journal of Scientific Exploration*, *15*, 299–329.

46. Jahn, R. G. (1989). Anomalies: Analysis and aesthetics. *Journal of Scientific Exploration*, *3*, 15–26.

47. Jahn, R. G. (1993). The complementarity of consciousness. In Rao, K. R. (Ed.), *Cultivating Consciousness for Enhancing Human Potential, Wellness, and Healing*. Westport, CT: Praeger (pp. 149–163). (Also available as Technical Note PEAR 91006, Princeton Engineering Anomalies Research, School of Engineering/Applied Science, Princeton University, Princeton, NJ.)

48. Dunne, B. J. (1997). Subjectivity and intuition in the scientific method. In Davis-Floyd, R., and Arvidson, S. (Eds.), *Intuition: The Inside Story*. New York: Routledge (pp. 121–128).

49. Jahn, R. G., & Dunne, B. J. (1997). Science of the subjective. *Journal of Scientific Exploration*, *11*, 201–224.