This article is a critical commentary on a whole area of research: *parapsychology*, a term coined by J. B. Rhine that covers phenomena such as *telepathy* the direct transmission of information from mind to mind. (Perhaps the field should have been called *paraphysics*, since it is supposed that accepted physical, rather than psychological, laws are being violated.) The landmark work is **Rhine, J. B. (1964) *Extra-sensory perception*.** (Boston: Bruce Humphries), and a flow of other publications by Rhine's associates and others. In this extended article, Lockhead shows how very small deviations from randomness in the to-be-guessed sequence can give rise to better- or worse-than-chance guessing performance.

Guess for Success: Sequential Behavior and Parapsychology

PsyCrit, January 2, 2007

ABSTRACT

People's guesses of successive random events are structured, which allows predicting their guesses with better than chance accuracy. Too, people perceive constrained sequences as random and perceive random sequences as constrained: Statistical randomness and psychological randomness are different. Guesses by a person reported to have psychic ability have the same structure as those of other people. Analyses of the parapsychological literature allow concluding that sequential structure explains all reports of psychic ability that are based on guessing sequential events.

Extra-sensory perception (ESP) got a major boost in the U. S. in 1934, when J. B. Rhine defined parapsychology as a scientific discipline. Rhine's research program was successful: many institutions became devoted to understanding how it is that some people seem to gain knowledge by means that elude the normal science of sensory systems. There is interest in ESP in many countries; it sells magazines, and is the basis for numerous science-fiction works, television programs and movies. As for scientific respectability, several journals are devoted primarily to parapsychology[1], and the Parapsychological Association became an affiliate of the American Association for the Advancement of Science in December 1969.

I became interested in this topic when students in my courses asked about parapsychology and about James B. Rhine, who made his reputation while he was a faculty member in my department at Duke. The result of my investigation is the following summary of critical aspects of the field as defined by Rhine, plus a reinterpretation of the data that caused the excitement.

The Operational Basis for Psi

Extra sensory perception, ESP, was proposed as an explanation of apparently non-chance guessing of random events, and people who demonstrate ESP are said to have *psi* ability. According to proponents of *psi*, four different phenomena have been demonstrated: clairvoyance, telepathy, precognition, and psychokinesis (Pratt, Rhine, Smith, Stuart, & Greenwood, 1940). However, arguments for these four phenomena are based on a single operation and so, as described ahead,

[1] For example, *International Journal of Parapsychology, Journal of the American Society for Psychical Research, Journal of Parapsychology, Journal of the Society for Psychical Research, Parapsychological Monographs, Proceedings of the Parapsychological Association.*

there is only one outcome to be explained. This is the seemingly above- or below-chance guessing of random sequences.

Clairvoyance is inferred when an observer correctly names an unseen event without any other person knowing what the object is. An example is when a person correctly predicts that Jeanne's photograph is in a randomly selected envelope. *Telepathy* is inferred when there is no physical object to be judged, but only a subjective image or thought to be apprehended from another person. An example is when a person correctly predicts that you are thinking about Jeanne. *Precognition* is inferred when an object is successfully predicted before it is selected by the second observer, as when Jeanne′s photo is predicted as being in the envelope before that envelope has been selected from a stack of possibilities. Finally, *psychokinesis* is inferred when a foretold event is later selected by some apparatus rather than by a person. An example is when the person correctly predicts that a machine will select Jeanne's name as the sweepstakes winner.

Because chance always allows a few correct guesses in such studies, Rhine usually conducted many trials to rule out the possibility of luck rather than psychic talent. Thus, typical ESP studies consist of many trials in which an observer predicts the contents of many successive envelopes, or other events. This method provides a sequence of guesses associated with a sequence of outcomes. ESP is said to be demonstrated when there are more (or fewer) correct predictions than expected on the basis of chance. Telepathy, clairvoyance, precognition, and psychokinesis are all based on such sequences of guesses followed by objects, or objects followed by guesses.

A common investigation of psi ability begins with the experimenter shuffling a pack of cards (for experimental convenience cards have been used most often, rather than envelopes or other events). Usually, an observer guesses the first card, which the dealer then turns over, guesses the second card, which is then turned over, then the third card, and so on throughout the pack. (Notice that this procedure gives *feedback* to the guesser.) This method produces a string of guesses paired with a string of cards. If the guesses are correct more often than expected by chance, it is inferred that the observer has psi ability. Whether the observer guesses cards or envelope contents or something else, and whether the cards are in the same room as the observer or are across the ocean, the data always consists of strings of guesses paired with objects. This is true for studies of clairvoyance and telepathy, where item selection precedes the guess, and also true for studies of precognition and psychokinesis, where item selection follows the prediction.

To operationalize studies of psi, Rhine created a pack of 25 playing cards, with five instances each of circles, squares, plus signs, star, and squiggly lines. He called these Zener cards, after his colleague Karl Zener, who suggested the method based on his psychophysical method of single stimuli. The idea is to shuffle the cards, and then ask people to predict each card before it is turned over. To minimize effects of luck on a single run, the cards are shuffled and guessed for many trial sequences. Because most reported demonstrations of ESP ability are based on data collected using these cards, this procedure is the focus of this manuscript.

It makes no sense to contest the existence of non-chance relations between item selections and observer guesses. That significant correlations have occurred is more than adequately documented (cf. Hansel, 1966). But it does make sense to ask about the interpretation of those correlations. The fact that what card will be turned over is guessed correctly more often than predicted by chance does not imply demonstration of psi or of anything paranormal. It simply means, for some reason, that there are situations in which predictions and events are related. Psi is a word, not an explanation. This paper suggests an explanation — these correlations are due to structured sequential behaviors by experimenters and observers.

Sequential guessing behavior

There are marked sequence effects in every study I have seen in which people guess a series of events. An example is an absolute-judgment experiment (Holland & Lockhead, 1968) where people judged which of ten different tones was presented on each of many trials. This was not an

ESP study; its purpose was to learn how well people identify sounds that differ only in amplitude. Tones were presented one at a time, with successive tones selected by a randomization algorithm, and the judges were provided feedback after each judgment

Performance was better than chance for all observers; 39% of all judgments were correct (chance was 10%). Thus, there were also many errors. These errors demonstrate two marked tendencies: People judged the value of each stimulus as overly like the one on the prior trial and as overly different from each of the six stimuli before them. This is *assimilation* followed by *contrast*: Judgments assimilate toward the value of the previous stimulus and contrast with several earlier stimuli.

Concerning ESP research, the same absolute-judgment study was repeated except for one change (Ward & Lockhead, 1971); the oscillator that produced the tones was turned off. In place of the tones, a fixed light was turned on to indicate when the tone would ordinarily have occurred. The observers' task was to judge what tone was selected even though none was presented. There was feedback as to the correct number after each response. This was a study of guessing behavior. A parapsychologist would call it a study of clairvoyance because the computer selected the random number before the subject guessed its value. Had the computer selected the number after the guess, it would be a study of precognition or psychokinesis; presenting the random number to the experimenter before the observer's guess would make it a study of telepathy.

There were 300 trials by each of 10 undergraduate students. The average of the 3,000 guesses across observers of the 10 equally probable events was 9.9+% correct; the chance expectation is 10%. No observer guessed the selected numbers with reliably better than chance performance, and thus none demonstrated psychic talent.

Although accuracy was at chance, there is considerable structure in the data. People tended to guess the value of each randomly selected stimulus as more like the just previous value (the feedback number [1 – 10] from the previous trial) than would occur if the guesses were independent of sequence. This is assimilation. They also tended to guess that this "stimulus" is less like earlier feedback than is expected by chance. This is contrast. For whatever reasons, people act as if they expect successive events to be relatively similar to one another, and as if they expect the current event to be relatively different from earlier events.

Sequence effects in a 'psychic'

Based on this result, it seemed interesting to ask if "normal" people and people reported to have psychic ability behave differently in sequential-guessing situations. If there are paranormal people who can guess random events successfully, then they must be using different information than the people reported above, whose guesses were not different from chance. Thus, such paranormals might behave differently than normals when guessing.

I looked for a person considered by parapsychologists to have marked psychic ability and who also believes he has such ability. The person we found was a native of Trinidad named Lalsingh Harribance, who had been brought to the Foundation for Research on the Nature of Man[2] in Durham, and then to the Psychical Research Foundation, two independent facilities in North Carolina, because of his reported psi talent. He was considered by some people to be a psychic, and he was described as a *sensitive* in the psychical literature[3]. This subject, in collaboration with a worker in the group supporting him, determined the particular study conducted. This is because he and his associate felt it important "for the conditions to be right" if Harribance's ability was to manifest itself.

[2] Now known as the Rhine Research Center.

[3] See the Proceedings of the Parapsychological Association, 1970. Several investigations with this person are reported as demonstrations of psi.

The study was conducted in 1973 when computers were novelties and these people were pleased to have the opportunity to demonstrate psi ability with a computer (Digital Equipment Corporation PDP8L). They were very cooperative and no payments were involved. Mr. Harribance reported that he preferred not to guess which of 10 numbers the computer had selected randomly, and he asked if a set of five things was acceptable instead. He also preferred to serve in blocks of 100 trials each, with performance accuracy presented in percent correct after each block, rather than in blocks of 300 trials. And he preferred to receive feedback as to the number selected by the computer immediately after each guess. The only requirement he insisted on was that the computer should select the random number before, not after, each guess. His preferences describe the procedure used.

There were four blocks of 100 trials. The psychic typed each guess, a number between 1 and 5, on a Teletype machine, which recorded the correct number beside this response ½-s later. The situation was self-paced and the average intertrial interval was about 9 s.

This observer guessed correctly on 85 of the 400 trials, for 21.1% correct. Chance expectation (binomial) is 80 plus or minus 8 correct. Thus, the observed value is less than one standard deviation from the expected value. Although the observer was pleased that he was correct on more than the expected 80 trials, this is chance performance. The observer talked rather often during the experiment (there were no instructions about this) and he gave varying reports of feeling that he was doing well and then badly and then well again throughout the study, but I found no correlation between his performance and these subjective estimates.

However, there are marked sequential dependencies in his data. Each guess depends on the feedback on the previous trials. A low feedback number tends to be followed by a low guess and a high feedback number tends to be followed by a high guess. As examples, when the feedback on the just prior trial was number 5, the mean guess for the next trial was 4.6; but when that feedback was number 1, the mean guess was 2.44. This is assimilation.

There also are effects of feedback two trials earlier, and these are reversed in sign from the one-trial back effect. When 5 was the feedback two trials ago, the mean guess on the current trial was 2.93, but when the feedback two trials ago was 1, the mean guess was about 3.37. This is contrast. Thus, this 'psychic' observer shows the same form of sequence effects as other observers, assimilation followed by contrast.

Predicting the Predictions

I tried to see if a simple sequential rule would predict the psychic's performance. In the language of parapsychology, this would be a study of precognition or telepathy, depending on whether I (i.e., my rule) guessed before or after the psychic subject generated his response. Rather than telepathy or clairvoyance, I assumed the psychic would behave essentially like other people, where guesses demonstrate assimilation followed by contrast. For simplicity, because a judgment was produced about every nine seconds (slow for usual psychophysical procedures, and likely to minimize the effects of stimuli several trials back), I only considered assimilation in making these predictions: When the feedback was 1, I predicted the next guess would be 1 or 2 and selected one of these numbers arbitrarily. Similarly, when the feedback was 2, I predicted the guess would be 1, 2 or 3 and selected one of these arbitrarily. Following 3, I selected from 2, 3, and 4, after 4, I selected from 3, 4, and 5, and after 5 I predicted the guess would be 4 or 5. This crude procedure predicted the observer's guesses on 35.4% of the trials (chance is 20%). Since the psychic observer did not predict the computer reliably, he did not show psi. But, since I predicted his guesses better than chance, did I?

Predicting what the observer will *not* guess, which was done after the study, provided a similar outcome. The rule for this: when #1 was the feedback, predict that the observer will not choose 3, 4 or 5; following 2, he will not select 4 or 5; after 3, he does not guess 1 or 5; after 4, 1 and 2 are avoided; and after 5, 1, 2 and 3 are not selected. By this analysis, 48% of the equally

likely five outcomes are predicted not to occur. Of the 396 trials (one trial is lost at the beginning of each of the four runs of 100 trials to get each sequence started), this analysis predicts incorrectly only 24 times (6.1%) rather than the expected 190 times (48%). Even stronger predictions are made with a more sophisticated model and by including contrast in the analysis, but these two examples make the point that the psychic's guesses were highly structured.

Randomization in the Psychical Literature

I am confident that I do not have psi, in spite of my guessing success. It is reasonable to ask, however, if the sequential information that allowed these significantly above-chance predictions might also account for reports of psi in the parapsychological literature. For this to occur, it is necessary only that stimulus sequences in those studies be constrained, i.e., non-random, in a particular way. This is because no guessing method can reliably produce non-chance choice behavior if there is no structure at all in the stimulus sequence. I therefore examined the psi literature to learn if stimulus selections there may be constrained.

Unfortunately, the question could be answered directly because stimulus sequences are not reported. Nonetheless, such information as is available strongly indicates that stimuli were not selected randomly in those studies where psi success is reported. Indeed, if ESP researchers follow the instructions of the leaders in their field, and there is every indication that they do, stimulus sequences are usually not fully random. As noted next, the instructions for conducting parapsychology experiments ensure both that stimulus sequences are structured and that observers have ample opportunity to learn about that structure.

A summary of 60 years of ESP research (Pratt, Rhine, Smith, Stuart, & Greenwood, 1940) concluded that "hand shuffling has been found by test to be adequate" (p. 65) to produce random stimulus selection. The only referenced support for this conclusion is a card-shuffling study using Zener cards (5 copies of each of 5 patterns) in 1934. That test is described as follows: "One 'middle spread' shuffle was given, in which a block of cards was removed from the middle of the pack and divided loosely between the top and bottom. Then one cut was made. The record was then taken, and the successive records were crosschecked. The average [number of matches] per 25 for the 1,000 trials was 4.7, or very close to the average of 5 to be expected (Rhine, 1934-35, p. 98)."

This means there was not a reliable average correlation of the serial position of a card type from one shuffle to the next. This may be so, but it does not test for sequential structure, which could have been substantial. This can be demonstrated by example: Consider a pack of alternating cards: 121212121212… Cut this deck after a l and correlate the resulting run with the original run. There are no commonalties; the same card type is never in the same serial position in the two sets, and, using Rhine's analysis, the correlation between the runs is a negative 1.0 or perfect contingency. Now, cut this second deck just as before, after a 1, and correlate this third string with the second string. This correlation is a positive 1.0, again complete contingency; both strings have a 2 in the first position, a 1 in the second position, a 2 in the third position, and so on. On successive random cuts, you should cut after a 1 about half of the time and after a 2 about half of the time. If you continue this procedure of randomly cutting and correlating successive decks, the number of matches will approximate the number of mismatches and the expected average correlation is zero. This satisfies the quoted criterion for randomness, which Rhine believed to imply no stimulus structure. But the constraints in the example sequence are so large that an alert observer would never make more than one error in each run (the first trial should be guessed wrong about half of the time) if feedback is given. Thus, this analysis for randomness in the ESP literature is wrong: The cards were not shown to be randomly presented.

Added shuffling may produce different outcomes, but large sequential constraints could still exist. One such example is that Pratt et al.'s (1940) proposed analysis for randomness cannot detect a sticky card, which a subject might readily detect. People are good at noticing contingencies, and

they might be expected to observe, for example, that B frequently follows A. This might have occurred often in that study but we cannot know because the sequences are not available. However, sticky cards are not needed to reject the method recommended by the ESP researchers. This is because their proposed cutting method essentially guarantees the same result, which is that some card adjacencies will be repeated across trials. The cards in many psi studies might be shuffled more completely than this recommendation, we cannot know this. But we do know it is not true that "hand shuffling [is] adequate." On the contrary, the procedure recommended for conducting ESP experiments is almost guaranteed to produce constrained sequences.

Seventeen years after this 1940 review, in a "one-volume summary of present knowledge about parapsychology" (Rhine & Pratt, 1957, p. v), the researchers again give instructions on basic psi test procedures. In discussing how to use the pack of 25 Zener cards, they say "it is best to develop a simple ritual of shuffling that insures this part of the procedure is not overlooked (p. 141)." This notion of ritual is not compatible with randomness. Yet, reassurance that all is well with the method is given again and again. Seven years later, on the occasion of republication of J. B. Rhine's 1934 book "that started all of the hullabaloo (1964, p. xxxiii)," Rhine takes the opportunity in a new appendix to reply to criticisms made on his reports over the intervening 30 years. The following quote is his summary of the importance of, and knowledge of, randomization in ESP research after 30 years of work. Rhine said

> One of the criticisms offered on the report is that the methods of shuffling the cards were not adequately described. This is a matter that may seem more important to the reader, since he is probably thinking about playing cards, than it actually is. In the first place, the cards used in the test for ESP have no individual numbering. There are 5 of each suit, and even if, by some sensory cue, the subject could locate, let us say, a circle, in order for this to indicate the cards following, he would need to know which of the five circles he had located [**1**, numbers in brackets refer to following comments]. If a single card is displaced and the pack is merely cut once, at a point unknown, then no one can know the order of the cards in the pack, even though he had memorized them all just before the cut [**2**]. Then, if we consider the work done at a distance, behind screens, with cards left in the unbroken pack until it is called down through, etc., we can see that such conditions eliminating all possible cues, even one shuffle of the simplest variety is quite enough to destroy knowledge of the order in the pack [**3**]. As state (sic) in the text, the pack of cards was always shuffled and cut. The pack being constituted as it is, it matters little how it is shuffled or how much. [**4**] (1964, p. 228).

Several assertions in this paragraph are problematic:

1. This is not correct. Given stochastic independence, the probability of any event following a particular event is the same as its probability given any other previous event. If it is known, for example, that a star has followed a circle, it does not matter which circle it was. Simply guessing star after each circle will produce better than chance performance because the conditional information increases the probability of a correct call. Additionally, because it is now less likely than 0.20 that a star follows some figure other than a circle, it is thus more likely than 0.20 that some other card type follows a circle. Accordingly, the probability of guessing correctly is further increased by never guessing star following other card types than circle.
2. This is simply wrong because it produces sequences that are constrained from trial to trial. Consider this example, which meets the instructions: Take any pack of cards in any order and label each card according to its serial position (or memorize the sequence). Then, in accord with the instructions, displace any card and cut the deck. The resulting sequence is highly predictable. If feedback is given after each guess and you guess according to the expectation that each card is followed by what occurred earlier, then you never need to make more than four errors in guessing all of the cards. Slightly more shuffling will produce less sequential constraint, but any situation with these randomization procedures produces sequential constraints.

3. This does not follow logically. There is no relation between card sequences and such factors as physical distance of the subject, or screening between him and the experimenter. This may be a core difficulty in this literature. Some researchers apparently believed that the basis for their findings lay in transmission systems, and thus other types of experimental factors, such as sequential structure, could be neglected or even ignored.
4. This asserts that one need not be concerned about these procedural matters. If researchers uncritically follow these directions, it is surprising that there are not more reports of psi than have been published.

In fairness while considering these criticisms, I should note that Rhine and his research group did not consider themselves specialists in statistics and employed statisticians to assure that their methods were correct. This was expensive and time consuming, and such care is evidence that Rhine and his co-workers were well-intentioned. Why their consultants did not detect the difficulties in their procedures I do not know. Possibly everyone involved thought the one test of randomness is all that is needed, even though an average zero correlation between stimulus strings is consistent with perfect sequential predictability within strings.

Magnitude of psi Effects

Although ritualistic shuffling produces recurring sequences, the effect of such a procedure, ineffectual as it is in producing complete randomness, might be small enough to be inconsequential. But it turns out that only a small effect is needed. The most popular demonstrations of psi show very, very small numerical deviations from chance. For example, the Pratt-Woodruff experiment (Pratt & Woodruff, 1939), particularly Part B, has been called "perhaps the most elaborately controlled (Rhine & Pratt, 1957, p. 47)" of the experiments that meet the criteria for a conclusive test of psi. That study involved 60,000 trials. Because of this large number of observations, less than one extra hit had to occur in each 100 trials, on average, to give statistically significant results; the reported result is that 0.816 extra hits per 100 trials were obtained. Taking this tiny numerical effect seriously requires us to be very certain indeed of the lack of structure in the stimulus sequence.

How good was the procedure in this famous study? For each run, the experimenter recorded on paper the sequence of 25 stimuli, recorded the sequence of 25 guesses, compared the two columns to calculate the number of hits, checked these data with the subject to assure accurate recordings (which also allowed the subject to study the structure of the stimulus string), wrote down the performance accuracy, filed the data, shuffled the pack of cards, gave a signal to begin the next run, and started the next run. All this occurred in an average of 2 minutes for each of the 2,400 runs. An experimenter who does all this and shuffles thoroughly for each run is amazing. And detecting less than one sticky card, or one recurring sequence of two cards, or making one recording error per four runs is sufficient to produce the positive findings in this "most elaborately controlled" psi study in the literature.

Feedback in the Psychical Literature

Information as to feedback provided in psychical studies is no more satisfying than that about randomization. Even though most studies do not report the feedback provided, a reasonable guess is that many studies provided feedback after each 25 trials, as described above, and it is known that some studies provided feedback on every trial. Because of this lack of complete information, the following conclusions are based more on published methods about how to conduct psi experiments than on the experiments themselves. If experimenters followed the directions, then feedback was ordinarily available.

Because of concern for "the psychological state of the subject," Rhine and Pratt (1957) say that the time between experimental runs is especially important and during this time the aim should be to "Recapture and restore the original zest for the experiment and thus counteract the effect of

the monotony on the subjects. The checking of the score sheets with the subject participating can be made a diverting procedure, especially by pointing out interesting effects and making optimistic interpretations." (p. 135) In describing a test procedure, they recommend during the experimenter's scoring: "Then follows the checking for hits, at which point it is a good plan to invite the subject to watch and insure accuracy." (p. 147)

Procedures to produce the "right psychological conditions" for psi thus are elaborately described, but terms like feedback, information, and knowledge-of-results do not appear in the indices or as a discussion topic in any of the major texts on parapsychology, at least not up to the time of this writing (1974), and essentially all of the experimental claims for psi were published before this date. Nonetheless, it is clear that feedback was given in many or all of the successful studies. Perhaps, asking the observer to study the sequence restores "the original zest" by giving him or her sequential information.

The possibility of relations between psi and feedback has been addressed in a few studies that report psi results, but the reports are not sufficiently definitive. For example, in a study "to assess the effects of feedback on subjects' abilities to discriminate between correct and incorrect ESP responses" (Honorton, 1970, p. 73), "no special precautions were taken since the feedback runs were employed solely as an experimental manipulation (p. 74)." Whatever this means, the paper does report that subjects given correct feedback later provided evidence of ESP, when the same test was repeated, while subjects given false feedback did not provide evidence in support of ESP. Concerning this, author suggested, "feedback facilitated the subjects' abilities to discriminate between correct and incorrect impressions" (p. 74). Perhaps it did this by providing information about the sequence. Such feedback could be useful because people can detect and remember and use information in sequences, and we know that they adjust their behavior on subsequent runs on the basis of that knowledge (Rose & Vitz, 1966).

With only one exception (Schmidt, 1969a,b, which has not been independently replicated[4]), I have not found a single report of psi where the method is reported sufficiently to evaluate the randomization procedure. On the other hand, none of the many studies using acceptable procedures that found evidence of psi. This agrees with the attempts of thousands of people over the ages to predict dice, cards, slot machines, the stock market, and many other factors that might improve their existence. The financial successes of casinos are ample evidence that people wish they could control the outcomes of the machines but cannot. A few people do guess profitably some times, and randomness requires some successes. But when such people try over and over again, the value of the casino stock increases.

Perhaps one ought not to be surprised that psychical researchers do not document randomization and feedback information. As in any field, experimenters report those factors considered necessary to allow someone else to replicate the experiment and these reported factors reflect the theoretical bias of the researcher. Thus, while investigating gravitational acceleration, physicists do not report the background music played while data are collected, physiologists do not usually report conversations while recording muscle potentials, and experimental psychologists historically did not generally report the time of day subjects provided data (although this is now known to be important in many cases). Likewise, the psychical researcher does not ordinarily report randomization procedures or the feedback given to the observer. All these omissions are for the same reason; music, conversations, time, and randomization and feedback are considered factors nonessential to the research.

[4] I have not attempted to criticize any studies beyond their use of randomization procedures. Many studies appear to be done carefully in other respects. Concerning Schmidt's report, hundreds of observers provided data in the same or similar device as he used and did not give evidence of psi; Schmidt could be a victim of chance processes.

Why Do People Believe in psi?

Why do many people accept the tenability of psi phenomena when "no subject has ever demonstrated his ability to obtain high scores when the test procedure is completely mechanized" (Hansel, 1966, p. 240), and when there is no well-defined theoretical basis for psi. 85% of the 1416 respondents to a questionnaire sent to readers of the *New Scientist* considered investigation of extrasensory perception to be a legitimate scientific undertaking (Evans, 1973). One possible reason has two factors:

First, it is important for people to attribute cause to events. Heider (l958a) argued that any perceived event gives rise to attribution as to its origin because such attributions are important to socialization and survival. "It makes a real difference, for example, whether a person discovers that the stick that struck him fell from a rotting tree or was hurled by an enemy (p. 16)." He extends Oppenheimer's consideration of "causal drive" as a basic drive, beside those drives for self-conservation and conservation of the species (Heider, l958b, p. 2).

When causes are correctly attributed to effects, we call it knowledge. We call it superstition when an assumed basis for causation is wrong. And if there is no apparent basis for events, causal agents are often invented — which may account for why inanimate things are treated as animate in many primitive societies, and why myths and multiple gods are invented to account for inexplicable events. Even in less-primitive societies, young children often treat clouds as animate. The assertion is that people will discover or create causes to attribute to events.

The second factor addresses why people accept assertions of psi based on correlations between nominally independent events. This may be because of a tendency to perceive that they correctly predicted random events more often than can be accounted for by chance. They believe that, more often than they can explain as coincidence, their preconceptions did occur, or they and another person had the same thought at the same time, or they began singing a tune another person was silently humming, or they predicted a lottery number. Parapsychologists accept the task of demonstrating that these correlations are due to parapsychological factors; to communication abilities that current science does not understand. Instead, they might accept the tasks of examining if the predicted events are in fact random and if the judge had any information concerning their structure.

Teraoka (1963) used the term "subjective randomness" when he showed constraints in observers' attempts to produce random strings of five things. His finding suggests that situations that appear independent may not be independent. This is consistent with Kahneman and Tversky′s (1969) concept of representativeness. They observed that people do not appropriately account for probability of occurrence when predicting outcomes.

The next section shows that random events are often perceived as constrained, and that constrained events are often perceived as random. Stochastic independence (mathematical randomness) and subjective randomness are not the same. I suggest in a moment that this discrepancy may account for many of the data that have seemed to support the field of parapsychology. The original interpretation of predicted random events that require explanation would be wrong. Rather, there may be constraints and thus predictability in occurrences that appear to be chance but are actually structured.

Comparing Subjective and Stochastic Randomness

Can people recognize randomness? John Monahan and I looked at this in a self-paced task in which the 23 students in an introductory statistics class each generated what they considered to be a string of 25 random numerals, using numbers ranging from 1-5. We then paired each of these 23 sets of 25 subjectively random numbers with a set of 25 stochastically independent numerals that also ranged from 1-5. This second set of numbers was generated by a computer-controlled algorithm that produces sequences of independent events that do not deviate from chance

expectancy according to any of several tests we conducted. Thus, there were 23 pairs of sets of numbers, with one set of each pair generated by a student and one set generated by a computer. Which set came first in each pair was randomly determined. The task of the experimental group, the 20 students in a second introductory statistics class, was to judge which sequence, in each of the 23 pairs of sequences, was the random one.

Each subject could have chosen anything between none to twenty-three of the stochastically independent sets as random. If the two sets looked identical to subjects, an average of 11.5 of each set should be chosen as random. If people cannot generate random events, and if other people can recognize what is random versus what is structured, more than 11.5 of the stochastically independent sets should be selected as random. However, on average, only 7.1 of the stochastically random sequences were selected as random, while 15.9 of the human generated strings were incorrectly selected as the random ones. 17 of the 20 subjects ($p < 0.001$) selected more human-generated sequences as random than computer-generated sequences, just the opposite of what is expected if people know independence when they see it.

Accordingly, subjectively random (human-generated) sequences are different from stochastically independent, random sequences, and subjectively random, structured, strings appear more random than unstructured strings. What appears random is often structured, and what appears structured is often random.

An Example of how Subjectively Random Sequences Could Produce psi Results

Imagine an experimenter who is testing a person's ability to guess the outcome of a tossed coin on each of several tosses. For data-recording convenience, the experimenter arranges two stimulus sequences ahead of time by flipping the coin 10 times. The outcome of one set of tosses is THTHTHTHTH ; the outcome of the other set is TTTTTHHHHH. Since he or she knows these are low probability events, the experimenter might be tempted to reject either sequence and toss again. This is particularly likely for well-intentioned experimenters sensitive to critics (cf. Hansel, 1966) who have expressed concern about possible fraud in the experiments. Suppose the experimenter does toss again, and now gets THHTHTTHTH. This is perfectly acceptable as a subjectively random outcome. However, the probability of this sequence is exactly the same as the probability of either rejected sequence. Stochastic randomness knows no structure or perceived form. The judgment that alternation or repetition does not occur randomly is a human notion, not a statistical one.

Should corrections similar to this ever occur, no matter how seldom, the experimenter's sincere attempt to conduct an honest and unbiased experiment by such occasional reshuffling the deck or re-tossing the coin can result in the subject performing different from chance. This is because the subject is no different from the experimenter. The observer also knows that the events to be guessed are random and "knows" that randomness means no perceived structure. If the sequence begins with THTH, the observer will probably not guess T on the next trial because of the "knowledge" that events are not structured. If the experimenter occasionally, and the subject frequently, operate with this concept, and if feedback is provided, performance will be better than chance. And, most importantly, because neither experimenter nor subject perceives any constraints in the stimulus sequence, it is easy for both to believe that a random set of events has been guessed with better-than-chance accuracy. Should events like these occur, and if people tend to infer causes to outcomes, it might be difficult for them avoid to a parapsychological conclusion.

This analysis also predicts another report of interest to the psychical literature. This is the statement that some observers show *negative psi* or *psi missing*, i.e., guessing events *less* often than they should by chance. It has frequently been suggested that some observers tend to cycle between psi and psi missing. Psi missing is of such interest that every issue of the *Proceedings of the*

Parapsychological Association, at least up to 1974 when I stopped searching, has several reports devoted to it.

To account for psi missing, consider an extreme example where a subjectively random sequence takes the form of avoiding stimulus repetition. Then, the sequence THTHTHTHTH is considered as more likely than TTTTTHHHHH (any non-alternating pattern will do). If the observer often guesses THTHTHTHTH, and if the experimenter occasionally constrains part of the sequence to relatively too-much alternation in agreement with the earlier argument, then both psi and psi missing will be reported on different occasions. When guesses happen to match the stimulus sequence, performance will be above chance, psi occurs. When the phase relation between the stimulus sequence and the guess sequence is off by one step, so the guesses are HTHTHTHT rather than THTHTHTH, performance will be below chance, negative psi occurs. Further, with reasonably short strings of data, relatively few occasions of no effect at all are expected. There should be apparent positive psi or negative psi much of the time. These constraints do not need to happen often to provide statistically significant effects. These are due to misperceptions of randomness and not to psychic ability.

This analysis also accounts for the common report that performance often seems to alternate from good to poor (from positive to negative psi) to good again, and for the frequent interpretations that that these presumed psi effects average out across the study to produce an overall result that only appears to be a null effect.

Notice also that this analysis does not require feedback. When the stimulus string is subjectively random, adding feedback only increases the likelihood that guesses and sequences will be positively correlated. Thus, one might expect more alleged occasions of negative psi when there is no feedback and of positive psi when feedback is given.

Conclusion

What has been called ESP is not several different things and there is no requirement to invoke telepathy, clairvoyance, psychokinesis, and precognition. Rather, the data sets that have been taken as evidence of these nominally different psychic talents are fundamentally identical: these are sequences of stimuli and guesses. Furthermore, the experimental method used for these studies creates sequential structure, which observers perceive and use, albeit unknowingly. Thus, it is not necessary to invoke psychic ability, for which there is no apparent mechanism, to account for the data. Sequential structure appears to account for all reported observations of ESP that are based on sequential guessing data.

Gregory R. Lockhead[5]
Department of Psychology and Neuroscience
Duke University, Durham, NC, 27708
greg@psych.duke.edu

REFERENCES

Evans, C. (1973) Parapsychology — what the questionnaire revealed. *New Scientist*, 57, 209.

Hansel, C. E. M. (1966) *ESP: A scientific evaluation.* New York: Scribners.

Heider, F. (1958) *The psychology of interpersonal relations*. New York: Wiley. (a)

[5] The original version of this critical commentary on parapsychology was prepared in 1974. I find no reason based on subsequent literature to modify these arguments.

Heider, F. (1958) Social perception and phenomenal causality. In R. Taguiri & L. Petrullo (Eds.*), Person perception and interpersonal behavior.* Stanford, Calif.: Stanford University Press. Pp. 1-21. (b)

Holland, M. K., & Lockhead, G. R. (1968) Sequential effects in absolute judgments of loudness. *Perception and Psychophysics*, 409-414.

Honorton, C. (1970) Effects of feedback on discrimination between correct and incorrect ESP responses. *Proceedings of the Parapsychological Association*, No. 7, 73-75.

Kahneman, D. S., & Tversky, A. (1972) Subjective probability: A judgment of representativeness. *Cognitive Psychology*, 3, 430-454.

Pratt, J. G., Rhine, J. B., Smith, B. M. , Stuart, C. E., & Greenwood, J.A. (1940) *Extra-sensory perception after sixty years*. Boston: Bruce Humphries.

Pratt, J. G., & Woodruff, J. L. (1939) Size of stimulus symbols in extra-sensory perception. *Journal of Parapsychology*, 3, 121-158.

Rhine, J. B. (1934-35) Telepathy and clairvoyance in the normal and trance states of a medium. *Character and Personality*, 3, 91-111.

Rhine, J. B. (1964) *Extra-sensory perception*. Boston: Bruce Humphries, 1964.

Rhine, J. B., & Pratt, J. G. (1957) *Parapsychology: Frontier science of the mind*. Springfield, Ill.: C. C. Thomas.

Rose, R. M., & Vitz, P. C. (1966) The role of runs in probability learning. *Journal of Experimental Psychology, 72,* 751-760.

Schmidt, H. (1969) Precognition of a quantum process. *Journal of Parapsychology, 33*, 99-108. (a)

Schmidt, H. (1969) Clairvoyance tests with a machine. *Journal of Parapsychology, 33*, 300-306. (b)

Teraoka, T. (1963) Some serial properties of *subjective randomness*. *Japanese Psychological Research, 5 (3),* 120-128.

Ward, L. M., & Lockhead, G. R. (1971) Response system processes in absolute judgment. *Perception and Psychophysics, 9*, 73-78.