A Short History of Psi Research

by

Robert Todd Carroll

Interest in the scientific study of psychical phenomena—as they were called in the 19th century—began with a movement known as *spiritism* or *spiritualism* in upstate New York in 1848 with the Fox Sisters: Kate, Margaretta, and Leah. Kate (age 12) and Margaretta (15) claimed to hear strange rapping noises in their bedroom. They convinced a few folks that they were getting messages from spirits. Soon they hit the road, managed by big sister Leah who was in her mid-30s. They went on tour performing séances, which became the rage in both the U.S. and Europe.

In the 19th century, several eminent scientists, including biologist Alfred Russell Wallace (1823-1913) and chemist Sir William Crookes (1832-1919), became interested in spirit communication. Both claimed that they had scientific demonstrations for the existence of psychic phenomena such as spirits *tilting tables*. In 1871, Crookes attended a Fox-girls séance in London and reported: "I have tested [the raps] in every way that I could devise, until there has been no escape from the conviction that they were true objective occurrences not produced by trickery or mechanical means" (Crookes 1874). In 1888, the sisters confessed that they had produced the raps by cracking their toe-joints and that they made bumping noises by fastening an apple to a string under their petticoats and surreptitiously bouncing it off the floor. This would not be the last time an eminent scientist was tricked by a subject of psychical research.

Crookes became interested in spiritualism after the death of his brother. He validated the levitation of Daniel “Dunglas” Home and the mediumship of Florence Cook, who was later proven to be a fraud and to have had an affair with Crookes. Cook and Crookes “had used the séances as a cover for their meetings” (Williams 2000: 66).

In 1853, physicist Michael Faraday (1791-1867) did his own experiments on table tilting and concluded that the phenomenon was due to “self-deception resulting from unconscious motor movements guided by expectation” (ideomotor action) (Hyman 1989: 85). The American chemist Robert Hare (1781-1858) at first agreed with Faraday but then did his own investigation. He developed an apparatus he called the Spiritoscope, designed to detect mediumistic fraud. In the process of testing his machine, he became a spiritualist convert.

Faraday hadn’t proven that *all* table tilting was due to ideomotor action. Some might still be due to spirits. And, of course, there could be dozens of ways a conjurer might produce table movements or the illusion of table movements. Faraday’s experiments might persuade those skeptical of spiritualism that the best explanation for the event was a physical and psychological one. But those who had experienced a powerful emotional upheaval at a séance would still maintain that something supernatural was occurring.

That *emotional* experience ignited an interest in testing psychic phenomena in several eminent scientists, including Wallace, Hare, and Crookes. Unfortunately, such an emotional investment easily fuels a pro-psychic bias that makes it difficult to do adequately controlled experiments on psychic phenomena.

It has been a common error made by eminent scientists in the history of parapsychology to believe that being intelligent, knowing how to conduct a scientific experiment, and being diligent against deception, cheating, or self-deception would be sufficient to guarantee a fair test of psychic powers. This became apparent during two of the first scientific experiments designed and supervised by several eminent men for the first society formed specifically to study psychic phenomena.

In 1882, Sir William Fletcher Barrett, a professor of physics at the Royal College of Science in Dublin, and a few friends, including the Cambridge philosopher Henry Sidgwick, formed the still-existing Society for Psychical Research (SPR). The goal of the society, in part, according to Sidgwick was to

> drive the objector into the position of being forced either to admit the phenomena as inexplicable, at least by him, or to accuse the investigators either of lying or cheating or of a blindness or forgetfulness incompatible with any intellectual condition except absolute idiocy.

SPR's first scientific study would have Sidgwick eating those words.

Barrett led SPR's first study (1882-1888). It involved a clergyman's four teenage daughters and a servant girl who claimed they could communicate telepathically. Barrett introduced a method for testing telepathy that was popular for more than a century, though it is rarely used anymore by scientific investigators: card guessing. He did a number of guessing experiments (of cards or names of persons or household objects) with the girls and came away declaring that the odds of their being able to guess correctly in one experiment "were over a million to one." The odds of their guessing correctly five cards in row were "over 142 million to one" and guessing correctly eight consecutive names in a row were "incalculably greater" (Christopher 1970: 10). More men of integrity with high degrees were brought in to witness the telepathic powers of the Creery girls and Jane Dean, their servant. All the scientists agreed that there was no trickery involved. How did they know? *They had looked very carefully for signs of it and couldn't find any!*

A skeptic might ask: *What are the odds that children can fool some very intelligent scientists for six years?* The answer is: the odds are very good. Almost immediately the scientists were criticized for being taken in by tricks amateurs could perform. It took six years for these rather intelligent men of the SPR to catch the girls cheating—using a verbal code—and discover their trickery. But that's not all. While one group of scientists was validating the Creery group, another from SRP was validating the amazing telepathic feats of a 19-year-old entertainer named George A. Smith and his partner in deception, Douglas Blackburn. Smith eventually became secretary of the SRP (Christopher 1970). Had Blackburn not eventually published a series of articles explaining how they fooled the scientists, the world might never have known the details of the trickery (Gardner 1992). The early scientific studies demonstrate the naïveté of the experimenters and the need for experts in non-verbal communication and deception, namely, conjurors or gamblers, to help them set up protocols to prevent cheating.

It took some time to sink in but eventually the experimenters realized that for some reason human beings like to deceive each other. They use all kinds of non-verbal signals to communicate, which can give the appearance of psychic transmission of information. They use glances (up, down, right, left for the four suits of a deck of cards, for example), coughs, sighs, yawns, and noises with their shoes. Other cheaters use Morse code with coins and various other tricks known to conjurers. Sometimes gestures to various parts of the body have a prearranged meaning.

Creery-girl and Smith-Blackburn stories are frequent in the literature on psi research but I will mention only one more: the Project Alpha fiasco. First, however, we'll examine Dean Radin's rather selective overview of the history of psi research presented in *The Conscious Universe*.

In 1889, Charles Richet, physiologist and Nobel laureate, experimented with hypnotizing a subject and having her guess the contents of sealed envelops. Radin says that the subject performed at odds far beyond chance. However, Christopher gives more details. The subject, Leonie B., identified 5 of 25 playing cards when tested in Paris by Richet. However, "when similar test with Leonie were repeated in London, her score dropped to pure chance average" (Christopher 1970: 18).

In 1885, the American branch of the Society for Psychical Research was established at Harvard University in Boston by Dr. Richard Hodgson (1855-1905), professor of legal studies at Cambridge University, and astronomer Simon

Newcomb. But the first American to publish a monograph (*Experiments in Psychical Research*) on his experiments with card guessing was John Coover.

Coover was Stanford University's first Fellow in Psychical Research. By 1917, he had done four large studies (trials of 10,000 or more) and reported that he had found nothing to support belief in ESP. The main experiment involved 100 pairs of subjects in 100 trials. Roughly half of these were for telepathy (experimental) and half were for clairvoyance (control). That is, in half the trials a sender looked at the card before trying to send a telepathic communication to a receiver. In the other half, the sender looked at the card *after* the receiver made his or her guess. Others examined Coover's data and found more than Coover did. Radin writes that the receivers' ability to guess the right cards rated 160 to 1 against chance (1997: 65). F.C.S. Schiller found the data showed odds greater than 50,000 to 1 against chance, but he used only the data from the fourteen highest-scoring subjects. Coover replied that he could find all kinds of interesting antichance events if he were selective in his use of the data (Hansel 1989: 28). In 1939, psychologist Robert Thouless found that if the data were lumped together from the main experiment, there were 44 more hits than expected by chance. Thouless suggested that the data supported some slight psychic effect. He calculated the odds of this happening by chance to be about 200 to 1. Coover attributed the excess hits to *recording errors* on the part of the experimenter (Hansel 1989: 26). Neither Schiller, Richet, nor Thouless, however, attempted to repeat Coover's experiment. That would have to wait until J. B. Rhine set up shop at Duke University. Radin says that Coover may have been more pessimistic about his data than others because of “disapproving pressure from his peers at Stanford” (1997: 65). However, Radin also notes that several studies have shown that a 1% error rate in recording is typical. Thus, Coover’s suspicion might well have been justified.

Richet was particularly vocal in his criticism of Coover’s work. Coover responded by proclaiming that it can’t be denied that fraud is frequent, general, and well known in psychical research. The witnessing of psychic phenomena by astute and eminent men, he said, has had a negative effect on the studies because it has led them to discount contrary interpretations of the same phenomena, ignore the lack of controls during those psychic experiences, and rely on the corroboratory testimony of others to such an extent that it has weakened the rigor with which the researcher should be expected to guard against fraud. Coover noted that in the other sciences the experimenter controls the conditions; but in testing psychical powers, the medium controls the conditions.*

While few remember John Coover, everybody knowledgeable of the history of psi research remembers Joseph Banks Rhine (1895-1980). In 1925, Rhine and his wife, Louisa, both with doctorates in biology (plant physiology) from the University of Chicago arrived at Harvard to study psychology, philosophy, and what Rhine would come to call “extra-sensory perception.” Both heard Sir Arthur Conan Doyle lecture on spiritualism and were impressed not only with his message but his serene demeanor. The possibility that spirits might be communicating with the living, said Rhine, was “the most exhilarating thought” he'd had in years. The Rhines sat in on a number of séances but were not completely taken in by their experiences. They were quick to claim that famed medium "Margery" (Mina, wife of Dr. Le Roi Goddard Crandon, a respected surgeon) was guilty of "brazen trickery." Yet, when they went to Duke in 1927 to work with William McDougall, their first investigation was of an alleged telepathic horse called Lady Wonder. They declared that they could detect no trickery and that the horse was genuinely telepathic. In a follow-up study, the horse couldn't perform and the Rhines declared that Lady Wonder had lost her psychic ability. A similarly clever horse had been studied by Oskar Pfungst in 1904 and it was found that the horse was responding to subtle visual cues. Had the Rhines been so inclined, they might have found the same thing with Lady Wonder. It turns out humans are as clever as horses and the phenomenon of unconsciously responding to sensory cues is now known as the *clever Hans phenomenon*. In any case, the Rhines took over the Duke lab from Dr. McDougall and ran it until Rhine's retirement in 1966. What did Rhine have after nearly forty years of scientific research on ESP and psychokinesis? He had a lot of data, a number of followers, but there was no Noble Prize on the horizon.

The Lady Wonder fiasco was just one of several blunders made by America's most preeminent name in parapsychology. His early results were similar to Coover's. He did a thousand trials of a card guessing experiment without finding any signs of ESP. He and Dr. Karl E. Zener did more experiments with numbers or letters of the alphabet sealed in opaque envelopes with the same non-results. Unlike Coover, however, Rhine did not give up. He and Zener changed the procedure to use what are now known as Zener or ESP cards, which gives the guesser a 1 in 5 chance of guessing a card correctly. They settled on a deck of 25 cards. Rhine believed that when someone was found who could do significantly better that 20% in guessing, that would be evidence for telepathy or clairvoyance. Some were so phenomenal (Adam J. Linzmayer, George Zirkle, Sara Ownbey, Hubert E. Pearce, Jr.), skeptics assume there must have been cheating. Rhine denied it. In any case, he described in detail the protocols and conditions under which his tests were made. Nobody thought Rhine was cheating but many thought he had been duped by his subjects several times. According to Milbourne Christopher “there are at least a dozen ways a subject who wished to cheat under the conditions Rhine described could deceive the investigator” (Christopher 1970: 24-25). Rhine did use a magician to observe one of his ESP phenoms, Hubert Pearce. When Wallace Lee (a.k.a. “Wallace the Magician”) was observing young Pearce, he performed at chance levels. Otherwise his scores were significantly higher.

Rather than admit that when controls are tightened it becomes more difficult to deceive investigators, Rhine and other psi researchers have often concluded that the controls have interfered with the paranormal realm. Some even claim that tight controls make the exercise of psychic power so difficult that it extinguishes it altogether in cases of severe scrutiny, such as when a trained expert in detecting deception is brought in. Experimenter control destroys trust and trust seems necessary for psychic powers to work, according to many psi researchers.

Rhine was undaunted by the criticism. In fact, he claimed in his first book (*Extra-Sensory Perception,* 1934) that he'd done over 90,000 trials and could justifiably conclude that ESP is "an actual and demonstrable occurrence." However, there were attempts to duplicate these trials at Princeton, Johns Hopkins, Colgate, Southern Methodist, and Brown, all without success. Critics could not find evidence in Rhine's report that he was as systematic and careful as one would expect a scientist to be making such an extraordinary claim. There was no evidence, for example, that Rhine realized how important it was to discuss how the cards were shuffled when doing the tests. He showed no awareness that the 1 in 5 odds that represent pure chance with the Zener deck could change if the cards were not perfect (which they weren't) and since certain strings of guesses would be ruled out with a universe of only twenty five entities. For example, no one would guess six or more circles in a row because the deck only contains 5, but in a truly random distribution of circles, 6 or more items of the same kind would be expected to come up occasionally. In fact, given the small size of the deck, the actual odds of guessing any given item might be different from the theoretical odds which are based on the assumption of extremely large numbers of trials where each item always has exactly the same chance of coming up. Even if verbal feedback is not given, which it often was, non-verbal signs might indicate to the subject that a guess was right or wrong and that would affect the next guess.

One indication that Rhine and his colleagues had little understanding of how theoretical statistics should be applied in the real world is revealed by their being puzzled how some subjects would do better than chance when they started off but their successes would taper off the longer they were tested. That is, the longer a successful subject was tested, the more his scores tended toward a chance distribution. Rather than take this as natural regression toward the mean (over time, all subjects should move toward chance if nothing paranormal is happening), Rhine, Radin, and some other parapsychologists explain it away by saying that it is due to the boring nature of the testing. They even have a name for it: *the decline effect*.

Radin goes through some of the criticisms made of the card experiments such as using hand shuffling instead of proper randomization procedures and the physical handling of the cards, which might allow the subject to read the card from impressions on the back of the card. He explains how it took some time before researchers realized that letting the subjects handle the cards or envelopes containing the cards opened the door to cheating. They first separated the experimenter and subject by a screen. Later they put them in separate rooms, and even in separate buildings to avoid the possibility of cheating or inadvertent communication by sensory cues.

But there were some things the researchers didn't seem to consider, such as the relationship of theoretical probabilities with real probabilities. In the 1930s, a magician by the name of John Mulholland asked Walter Pitkin of Columbia University how does one determine the odds against matching pairs with five possible objects. Of course, Mulholland didn't have a computer to do his dirty work for him, so he printed up 200,000 cards, half red and half blue, with 40,000 of each of the five ESP card symbols. The cards were mechanically shuffled and read by a machine. The result was two lists of 100,000 randomly selected symbols. One list would represent chance distribution of the symbols and the other would represent chance guessing of the symbols. How did they match up? Well, they didn't. The actual matches and what would be predicted by accepted theoretical odds didn't match up. The total number was 2% under mathematical expectancy. Runs of 5 matching pairs were 25% under and runs of 7 were 59% greater than mathematical expectancy. The point is not whether these runs are typical in a real world of real randomness or whether they represent some peculiarity of the shuffling machine or some other quirk. The point is that Rhine assumed that statistical probability, which assumes true randomness and a very large number of instances, applies without further consideration to decks of 25 cards shuffled who knows how or how often.

Rhine and all other psi researchers have assumed that a*ny significant departure from the laws of chance is evidence of something paranormal.* While cheating should be of concern to paranormal investigators, there should be more concern with this assumption. There are two problems with it, one logical and one methodological. The assumption either begs the question (assumes what needs proving, namely that deviation from chance is evidence of psi) or commits the fallacy of affirming the consequent (*If it's psi, then the data will deviate from chance. The data deviate from chance. So, it's psi.*). The assumption is also questionable on methodological grounds. Studies have shown that even when no subjects are used there is significant departure from what would be expected theoretically by chance (Alcock 1981: 159). For example, Harvie "selected 50,000 digits from various sources of random numbers and used them to represent "target cards" in an ESP experiment. Instead of having subjects make guesses, a series of 50,000 random numbers were produced by a computer." He found a hit rate that was significantly less than what would be predicted by chance "If such significant variation can be produced by comparing random strings with random strings, then the assumption that any significant variation from chance is due to psi seems untenable (Alcock 1981: 158-159).

In any case, it seems to be a bit of an exaggeration for Radin to claim that statistician Burton Camp "finally settled" the issue of the statistical criticisms when he declared that Rhine's "statistical analysis is essentially valid" (1997: 95-96).

Another example of Rhine's lack of sophistication with probabilities comes from the fact that when he found subjects who scored consistently *below* chance, he did not see that this would be expected by the laws of chance. Instead, he took this to be evidence of psychic phenomena. He claimed that subjects who didn't like him would consciously guess wrong to spite him (Park 2000: 42). Some parapsychologists accept this explanation and the phenomenon is termed *psi-missing*.

Rhine did not convince the scientific community of the reality of ESP, despite his claims that his subjects had been "carefully witnessed" and that he had put into place "special conditions" that "completely eliminates all chance for deception." That was about as much detail as he gave the world. It wasn't enough. His lack of detailed documentation simply added to the perception of many skeptics that ESP researchers are too trusting and careless in setting up their protocols.

Rhine also did many PK experiments with dice—in which subjects try to will the outcome of the roll of a die—beginning in 1935. He didn't publish anything on the subject until 1943, however. Many such experiments were done in several labs between 1935 and1989. In 1986, Radin and Diane Ferrari did a meta-analysis of the dice experiments data and found that the control studies yielded 50.02% (odds against chance of 2 to 1), while the experimental studies yielded 51.2% (odds against chance of a billion to 1.) A meta-analysis takes the data from many individual studies and analyzes the data as if it were produced by a single large study. The validity of a meta-analysis depends on two important, but conflicting, factors: (1) there should be no selective reporting (using only *some* of the studies that have been done) and (2) one should use only studies that are done with proper protocols and controls. Suffice it to say that when Radin lumps together the data from 142 articles published between 1880 and 1940 and claims that they represent 3.6 million individual trials by 4,600 subjects in 185 experiments, he is not doing something that is clearly justified. He does not seem to be justified in claiming that the hit rate was significantly over the 20% chance rate and "sufficient to settle the question about the existence of psi perception." Some of the studies are of questionable value and there is no way these studies are all of equal value. If you add up 185 experiments, many of which use questionable protocols, you don't get one big unquestionable result.

Of the dice experiments, Radin says that he and Ferrari took the information from the studies and "for each study we calculated a 50-percent equivalent chance hit rate" (1997: 134), but he doesn't say how this calculation was done. (Note: only 7 has a 50% chance of being rolled; all other combinations are less than 50%, ranging from 1/12 (8.3%) for 2 and 12 to 5/12 (41.6%) for 6 and 8.)

Radin notes that other analyses showed that the results were not due to a few investigators or studies nor to the file-drawer effect, though the latter remains a problem. (The file-drawer effect refers to the practice of not reporting studies that get negative results.) He doesn't say how he calculated that one would need 17.974 studies in the drawer per published study to nullify the data.

The most interesting thing Radin did, in my opinion, was to correct for dice bias. He tested and supported the hypothesis that the more dots on a die face the less mass and the less mass the more likely it is to come up on top (1997: 137). (This hypothesis was validated except for 3 dots, which didn't seem to fit the pattern.) But, even correcting for dice bias, he had 69 experiments that followed "balanced-protocol" criteria—die faces were equally distributed among the six targets. He claims that they still got better than chance results but he doesn't specify how much greater, though he says that the odds of getting his results against chance were *more than a trillion to one*.

Thus, he says, neither chance, nor the quality of the studies, nor selective reporting can explain away the data.

On the other hand, the dice experiments were critically evaluated by Edward Girden of Brooklyn College. Radin makes an oblique reference to Girden's work by footnoting him, along with G. Murphy's report on a Girden paper on psychokinesis, viz., footnote 23 on page 133, which reads "By 1989 dice experiments had been reviewed an criticized numerous times over the years, but in spite of all the experiments and review, no clear consensus had emerged." This seems to be Radin's way of admitting that not everybody agreed with his rosy analysis, but he doesn't go into detail regarding Girden's concern. C.E.M. Hansel does (1989). "Only one of the early experiments [1934-1946] employed a control series" and this experiment "provided no evidence for psychokinesis but clear evidence for bias of the dice, since the dice tended to fall with the 6 face uppermost, whether it was being wished for or not" (Hansel 1989: 172). Among the later investigations, out of thirty studies thirteen were positive and the rest didn't produce above-chance scores (Hansel 1989: 174). Girden also applied criteria that Rhine and Pratt (*Parapsychology* 1954) had said were conditions for a conclusive PK test—*having two experimenters, true randomization of targets, and independent recording of targets, hits, and misses*—and on these criteria "none of the thirteen tests giving positive evidence for psychokinesis can be regarded as conclusive, whereas several of the remaining seventeen investigations that failed to provide such evidence do satisfy the requirements" (Hansel 1989: 174).

While Radin gives extensive coverage to many researchers, he ignores the work of British mathematician S. G. Soal (1889-1975). Soal was a critic of Rhine's protocols and claimed that he was going to improve on Rhine's methods and systematically exclude sensory leakage or deception of any kind in his experiments. By 1939 Soal had tested over 160 subjects with more than 128,000 card guesses. He said that he had found no evidence of telepathy, i.e., *nothing of statistical interest*. Then, it was suggested to him by Whately Carrington that he go data mining for *displacement*. Soal claimed to find statistically significant results (i.e., not likely due to chance) with two of his 160 subjects when he correlated guesses with cards *preceding* or *following* the target cards. He and others took this as evidence of clairvoyance. We now know that Soal didn't just go data mining. He went data *changing* (Alcock 1981: 140-141). Yet, it was Soal who said of Rhine's claims of verifying a telepathic horse and of being hoodwinked by Pearce that in Britain such stunts would be "quickly exposed as frauds or conjuring tricks." In the U.S., with its lax standards for scientific investigation into psychic matters, "they are proclaimed genius with a blare of trumpets" (quoted in Christopher, p. 29). Whatever Soal may have believed about the integrity of testing and the experimenter when he began his research, his methods seem to have been totally compromised by the time he had finished his best work. In 1954, he published a report on his experiments and seemed to brag that he hadn't bothered with "ultra-rigorous precautions on fraud" because, he said, if the experimenters "are not to be trusted, then there is no point whatever in their doing experiments" (quoted in Christopher, p. 30). However, the

Soal-Goldney experiments, (1941-1943) which were intended to be a replication of the precognitive abilities of Basil Shackleton, turned out to be a replication of dishonesty by a scientist.

In one sitting, Shackleton's success at guessing one card ahead was so great that Soal calculated the odds against chance to be greater than 10^{35} to 1. In another, the odds against chance were calculated to be 10^{11} to 1. There were several other sessions in which Shackleton performed at phenomenal levels when measured against chance expectation. When a procedure was introduced that sped up the process of testing, Shackleton performed significantly above chance levels at guessing the card two ahead of the target. It looked as if parapsychology had solid scientific proof of psychic ability.

The Soal-Goldney experiment was hailed by many as an example “of the strength of evidence for the reality of ESP” (Robert Thouless). The philosopher C. D. Broad wrote: “Dr. Soal's results are outstanding. The precautions taken to prevent deliberate fraud or the unwitting conveyance of information by normal means are described in great detail, and seem to be absolutely water-tight” (quoted in Hansel 1989, p. 106). G. Evelyn Hutchinson, a biology professor at Yale University wrote: “Soal's work was conducted with every precaution that it was possible to devise” (quoted in Hansel 1989, p. 106). J. B. Rhine compared Soal's work favorably to his own.

Today, it is generally recognized that Soal altered and faked the data, probably unbeknownst to Mrs. Goldney. His fraud did significant damage to parapsychology (Alcock 1981: 140-141). It was now apparent that one had to protect against cheating not only from the subjects but from the experimenters as well.

By the second half of the twentieth century, protocols in psi research had become much more sophisticated than in its early years. Advances in technology would significantly reduce some of the earlier problems with data recording, randomization, sensory leakage, and so on. In the 1960s, physicist Helmut Schmidt started using random event generators to do micro-PK (MPK) experiments. According to Radin, over the years Schmidt provided solid scientific support for the PK hypothesis and the people involved in the PEAR group replicated Schmidt's work in 258 experimental studies and 127 control studies. C.E.M. Hansel, however, claims that regarding all the studies done after 1969 and before 1987 that attempted to replicate Schmidt's work: "The main fact that emerges from this data is that 71 experiments gave a result supporting Schmidt's findings and 261 experiments failed to do so" (Hansel 1989: 185). Radin says that between 1959 and 1987 there were 832 RNG studies by 68 investigators: 597 experimental studies and 235 control studies. The best of these studies are those done by Robert Jahn and his group at the Princeton Engineering Anomalies Research laboratory (PEAR), which closed down at the end of February, 2007.

From 1966-1972, there were a number of dream telepathy experiments at Maimonides Medical Center in Brooklyn, New York, conducted by Montague Ullman and Stanley Krippner. Skeptics criticized these studies for a variety of reasons. Radin mentions none of the skeptical critiques, which include data on attempts at replication that failed when controls got tougher (Hansel 1989: 243-254).

Radin provides a telling anecdote that reveals a problem with much psi research: the problem of clearly identifying before the research begins exactly what will count as information transfer. He mentions two receivers in the dream telepathy experiments, which involved a sender concentrating on a target (Max Beckmann's painting *Descent from the Cross,* which depicts Christ being taken down from the cross) and a receiver whose dreams are supposed to be influenced by the sender. The sender in this case was also given some additional visual aids to work with: a crucifix, a Jesus doll, nails, and a red marker. He was given instructions to nail the doll to the crucifix and use the marker to color the body as if with blood. One receiver dreamed of a speech by Winston Churchill and the other of a "native ceremonial sacrifice." There was a reference to "sacrificing two victims," something about "destroying the civilized," and "the awe of god idea." Radin comments on the symbolic significance of "church-hill." There is nothing in either dream of the crucifixion in all its gory representation, yet these dreams are considered successful telepathic "hits." After all, Christ died on a hill, a church is named after him, the crucifixion is looked at as a sacrifice by Christians, Christ is both man and God (*two* victims?), and there was a god of some sort mentioned in one of the dreams. However, there are literally thousands of items that one might retrofit to these dreams. By allowing loose and symbolic connections to be made after the fact in order to evaluate the accuracy of the “telepathy” may be more a measure of the *cleverness* and *desires* of the judges than of the paranormal powers of the participants.

Thus, it is probably not worthwhile to evaluate Radin’s meta-analysis of some 450 dream studies. I agree, however, with Radin’s comment on these studies: “All we know from the present overview is that chance can be soundly rejected as one of many possible explanations for the results observed in these studies” (1997: 73).

From 1972-1994, there were a number of remote viewing experiments, primarily at the Stanford Research Institute (SRI), which was “a scientific think tank affiliated with Stanford University” until the late 1970s when it became the independent SRI international (Radin 1997: 98). In 1972, physicists Harold Puthoff and Russell Targ founded the SRI remote viewing program. Targ left in 1982; Puthoff left in 1985 (Marks 2000: 71). Physicist Edwin May joined SRI in 1975 and became the director of the program when Puthoff left.

In 1990 the program moved to another “think tank,” Science Applications International Corporation (SAIC), a major defense contractor and a Fortune 500 company with some 38,000 employees worldwide (Marks: 73).

Radin says the RV program “finally wound down in 1994.” He doesn’t mention that the CIA shut it down because they were convinced that after 24 years of experiments it was clear that remote viewing was of no practical value to the intelligence community (Marks: 75). The CIA report noted that in the case of remote viewing there was a large amount of irrelevant, erroneous information that was provided and there was little agreement observed among the reports of the remote viewers (Marks: 77). Radin doesn't mention that May objected to the CIA report because it didn't make note of the fact that he had four independent replications of remote viewing. May didn't publicize the fact, however, that there were also at least six reported instances of failed replication.

Radin makes it sound like the government's money was well spent (somewhere between 20 and 24 million dollars over more than 20 years). It's easy to understand why remote viewing would be of interest to the military and spy agencies. But it is difficult to understand why those agencies would abandon RV if it was as successful as Radin makes out.

Radin doesn't evaluate the studies. Rather he pulls out some selective examples of successes, i.e., reports or drawings that were judged to be very accurate. What he doesn't reveal is that one of the major flaws in all the later RV studies—done under the direction of May—which were better designed and controlled than the ones done by Targ and Puthoff, were fatally flawed because May, the director of the program, was the sole judge of the accuracy of the reports and he conducted the experiments in secret (which made peer review and replication impossible). David Marks writes that he tried for years to get May to let him look at his data, but May wouldn't allow it (Marks 2000).

There were hundreds, maybe thousands of trials, where a remote viewer would draw something and give a verbal report of what he was seeing. It would be highly unusual if there weren't some that would seem very accurate for the targets. Since it was never required for success that the drawing or report be exact, it is always possible that an ambiguous image will be seen as fitting a particular target especially if the judge knows what the target is. Furthermore, we have only May's word for it that the very detailed descriptions were spot on as he says they were. Nobody's ever seen his all his data.

Radin is correct that all possible paths for sensory leakage can be controlled for in RV experiments but he doesn't mention the actual method used by May to judge the results. Radin notes that "a judge who was blind to the true target looked at the viewer's response (a sketch and a paragraph or two of verbal description) along with photographs or videos of five possible targets. Four of these targets were decoys and one was the real target" (1997: 100). In fact, when this protocol was used by Marks he was unable to replicate either the RV experiments of Targ and Putoff or those of May. An analysis of the Targ and Puthoff experiments was done by Marks and he found that they systematically violated this rule about blind judging. Marks found substantial evidence that Targ and Puthoff cued their judges by including dates and references to previous experiments in the transcripts, "enabling the judges to successfully match the transcripts against the list of target sites" (Marks: 57). There were a number of other flaws in the Tare and Puthoff experiments detailed by Marks (2000: chapter 3) and Randi (1982: chapter 7), none of which are mentioned by Radin.

Radin makes it sound like constructive criticisms led researchers to refine their techniques to prevent any cheating or inadvertent cuing, but nothing could be further from the truth. He is correct that May's positive results of his analysis of all the RV studies done at SRI can't be explained by chance. But he's wrong to claim that "design problems couldn't completely explain away the results" (1997:

101). The SRI studies were fatally flawed and could not be replicated (Marks 2000). The SAIC studies (1989-1993) were likewise flawed.

Radin's account of the CIA commissioned report is also incomplete. It's true that Jessica Utts and Ray Hyman were the evaluators of the SAIC studies. Utts coauthored several papers with Ed May, so she was not a disinterested party and Hyman is a known skeptic, so he's not disinterested either. But the CIA wanted a review done quickly and had to pick people knowledgeable of the studies and they wanted a believer and a skeptic, for balance I suppose. They were to focus on two issues: 1. Is there scientific justification for the reality of remote viewing? 2. Is remote viewing of practical use for intelligence gathering? Utts claimed there was good statistical evidence to support the reality of RV; Hyman disagreed. Marks also disagrees, mainly because only one judge was used throughout the experiments and he was the principal investigator.

> …given the Principal Investigator's familiarity with the viewers, the target set, and the experimental procedures, it is possible that subtle, unintentional factors may have influenced the results obtained in these studies. (Marks: 76)

The report concluded that remote viewing is of little value and the CIA terminated the program known as STAR GATE.

Radin describes the SAIC studies as "rigorously controlled sets of experiments that had been supervised by a distinguished oversight committee of experts from a variety of scientific disciplines" (1997: 101). But he makes no mention of the fact that May alone judged all the cases and has not let anyone see all the data, even though it is all unclassified. And even though the SRI studies were fatally flawed, the SAIC folks and most believers in psi consider them excellent studies that have proven RV.

Radin is not quite accurate when he says the "government review committee" came to six general conclusions. His reference is to Jessica Utt's article, "An assessment of the evidence for psychic functioning" in the *Journal of Scientific Exploration*. Utts did not represent the government. In any case, the first item she listed was that free-response remote viewing was more successful than forced-choice remote viewing. This hardly seems like a major discovery. 2. Some people performed better than others. 3. Only about 1% of those tested were very good at remote viewing. 4. Training is worthless and RV ability can't be improved. 5. Feedback seems to enhance performance. 6. Shielding the target made no difference to the quality of RV.

So, Utts, who is an active researcher in the field, reports that the evidence is in and it's been replicated. We don't need to look for proof any longer. Whereas Hyman, whom Radin calls "the devil's advocate" for some reason, agreed that the effect sizes in the SAIC studies aren't likely due to chance, file drawer effect, or inappropriate statistical testing or inferences. The SAIC studies were well designed, he says, but remember Hyman did not have access to the data nor did he apparently know that May was the only judge in those studies.

Radin mentions that Julie Milton did an analysis of 78 free-response psi experiments published between 1964 and 1993 and found that "the overall effect resulted in odds against chance of ten million to one" (1997: 106). But he doesn't mention that only two of the studies had proper safeguards for the crucial protocol of “avoiding giving cues to judges and keeping the experimenter blind to the identity of the target in telepathy and clairvoyance” (Marks: 93). Nor does Radin mention that 26% of the studies failed to provide adequate safeguards regarding the person transcribing the subject’s descriptions being blind to the target’s identity and that this was associated with a significantly higher effect size than the studies that contained this safeguard (Marks: 93-94). Marks reminds us that “*statistical significance and real-world importance are not the same thing*” (2000: 94).

From the mid-1970s to the mid-1980s, some of the best of the scientific studies on telepathy were done: the ganzfeld (*whole field*) experiments conducted by Charles Honorton, William Braud, and Adrian Parker. These experiments begin with the assumption that certain mental states are more conducive to psi. In particular, they believe that the meditative state, the dream state, the hypnagogic state, the hypnotic state, a sensory deprivation state, and certain drug-induced states are conducive to psi. These states, it is believed, have in common “reduced sensory input.” It is thought that the mind in this state is alert and receptive. [For more on the ganzfeld experiments, see http://www.skepdic.com/ganzfeld.html.]

Radin seems to believe that the filtering of sensory data that is a function of ordinary consciousness may also be filtering out extrasensory data. But the really interesting thing about these experiments is that skeptics (Ray Hyman 1989: 20-75) had input in creating the protocols. As Radin notes: “Most of the ganzfeld experiments took advantage of lessons learned in past psi research, thereby avoiding many of the design problems discovered by early experimenters" (1997: 74).

From 1978-1987 at the Princeton Engineering Anomalies Research Laboratory, there were studies on **precognitive remote viewing** (PRP). In these RV studies, the receiver reports his impressions *before* the sender chooses a target. The PEAR folks not only claim many successes but Radin reports that they figured out a way to calculate the odds against chance of such activity and that their overall data were 100 billion to 1 against chance. Radin gives one example of how successful these studies were and the example indicates the problem with allowing a vague or ambiguous stimulus to be described and then later a judge decides whether there is a fit. The RVer describes being *inside* a large bowl. The target selected later was a radio telescope, which, according to Radin, "resembles a large bowl." With loose judging standards such as these there is no need to look for other explanations as to why they were able to succeed with odds significantly higher than chance.

In 1986, Robert Jahn, Brenda Dunne, and Roger Nelson of PEAR reported on millions of trials by 33 people over seven years trying to use their minds to override random number generators (micro-PK). In 1987, Radin and Nelson did a meta-analysis of several hundred experiments involving RNGs and found that they produced odds against chance beyond a trillion to one. The PEAR studies are considered by critics to be the best designed and controlled experiments on PK. We will return to these studies later. [For more on the PEAR studies see http://www.skepdic.com/pear.html.]

From 1979-1983, Peter Phillips did experiments on psychokinesis. However, he was the victim of a hoax, code-named Project Alpha. The hoax involved James Randi, Steve Shaw (a.k.a. Banachek), and Mike Edwards. Randi trained two young mentalists/magicians—Banachek was 18 and Edwards 17 when the project began—to fake psychic powers while being investigated in a serious scientific setting. They were able to fool the scientists for four years through more than 160 hours of experiments on their paranormal powers.

In 1979, James S. McDonnell, board chairman of McDonnell-Douglas Aircraft and devotee of the paranormal, gave $500,000 to Washington University in St. Louis, Missouri, for the establishment of the McDonnell Laboratory for Psychical Research. Randi saw this as an opportunity to disprove the complaint of many parapsychologists that they were unable to do properly controlled experiments because of lack of funding.

Randi believed that funding was the least of their problems. In his view, the main obstacle to parapsychology was its “strong pro-psychic bias.” This bias blinds researchers to numerous flaws in their protocols, almost all of which are related to their naiveté regarding human deception and their inexperience at detecting such deception. Some parapsychologists, such as Stanley Krippner, then president of the Parapsychological Association, agreed with Randi that qualified, experienced conjurors were essential for design, implementation, and evaluation of experiments in parapsychology, especially where deception—

involuntary or deliberate—by subjects or experimenters, might be possible. But many parapsychologists ignored Randi, as they had been ignoring similar criticism for more than a century.

Randi trained Banachek and Edwards so well that out of 300 applicants they alone were selected as subjects. The director of the McDonnell Lab was physics professor Peter R. Phillips, who had been dabbling in parapsychology for about ten years. He told the press that his lab would investigate "psychokinetic metal bending (PKMB) by children." Randi sent Phillips a list of protocols (he called them "caveats") that should be instituted when testing human subjects to prevent deception. One of the things he warned him about was not to allow the subjects to run the experiments by changing the protocols, a practice Randi knew is a common ploy of alleged psychics. He also warned that capricious demands by subjects might well be the means of introducing conditions that would permit subterfuge. Randi also advised that a conjuror be present during the experiments and even volunteered himself at his own expense to do the observing. Phillips told Randi he was quite confident he could conduct proper experiments without Randi's help. Randi writes

> Though I had specifically warned Phillips against allowing more than one test object (spoon or key, for example) to be placed before a subject during tests, the lab table was habitually littered with objects. The specimens were not permanently marked, but instead bore paper tags attached with string loops. Banachek and Edwards found it easy to switch tags after the objects had been accurately measured, thus producing the

> illusion that an object handled in the most casual fashion had undergone a deformation (Randi: 1983a).

Phillips and his lab assistants became convinced the boys had psychic powers but they also thought of their work as exploratory. In 1981, they took a videotape of the Banacheck/Edwards sessions to a convention of the Parapsychological Association. Their colleagues at the convention are said to have laughed at the video and noted numerous weak spots in their protocols.

Soon afterward the McDonnell folks began instituting protocols that had been suggested by Randi. Almost simultaneously they found that the boys seemed to have lost their ability to produce psychic effects. It was at this point that the boys were dismissed and Randi made the hoax public. Randi's take on the project after it was completed was

> If Project Alpha resulted in Parapsychologists (real parapsychologists!) awakening to the fact that they are able to be deceived, either by subjects or themselves, as a result of their convictions and their lack of expertise in the arts of deception, then it has served its purpose. Those who fell into the trap invited that fate; those who pulled back from the brink deserve our applause (Randi: 1983b).

Twenty years later Randi observed that "the effect of Alpha didn't last long" (personal correspondence). This exposé, like many others before it, has had little impact on the parapsychological community. Rather than thank skeptics for vividly demonstrating how easy it is for very intelligent, highly trained professionals to be fooled by conjurers, they ignore the skeptics. Or worse, they accuse them of "offensive incredulity."

In 1994, biologist Rupert Sheldrake published a report on psychic dog, Jaytee, a terrier who has precognition (*Dogs That Know When Their Owners Are Coming Home: And Other Unexplained Powers of Animals*). In 1998, psychologists Richard Wiseman and Matthew Smith tried to replicate the Jaytee experiment and failed. Sheldrake also published a report on a psychic parrot. [For more on the psychic parrot work see http://www.skepdic.com/nkisi.html.]

In 2002, psychologist Gary Schwartz published *The Afterlife Experiments* about his research into the ability of mediums to get messages from spirits. Schwartz, too, is claiming that he has tested mediums and that their performances have exceeded all odds against chance. [For more on Schwartz's work see http://www.skepdic.com/essays/gsandsv.html.]

Not everyone agrees, however, that a review of the literature reveals odds against chance in the overall database that are on the order of "a billion trillion to one" (Radin 1997: 97). But even if they are, not everyone agrees that such deviation from the laws of chance support the psi hypotheses. As Milbourne Christopher put it:

> Many brilliant men have investigated the paranormal but they have yet to find a single person who can, without trickery, send or receive even a three-letter word under test conditions (Christopher 1970: 37).

Nor have we yet to find a single person who can move a pencil across a table without trickery or without touching it.

Finally, Radin accepts Gertrude Schmeidler's notion of the "sheep-goat" effect, that believers get good results and skeptics get negative results in psi experiments. He reminds us: "Together, culture, experience, and beliefs are potent shapers of our sense of reality. They are, in effect, hidden persuaders, powerful reinforcers of our sense of what is real" (1997: 108).

§

last updated March 7, 2007

Bibliography

Alcock, James E. (1981). *Parapsychology: Science or Magic.* Pergamon Press.

Christopher, Milbourne. (1970). *ESP, Seers & Psychics.* Thomas Y. Crowell Co.

Crookes, William. (1874). "Notes of an Enquiry into the Phenomena called Spiritual during the years 1870-1873." *Quarterly Journal of Science*, January. http://www.survivalafterdeath.org/books/crookes/researches/notes.htm

Hansel, C. E. M. (1989). *The Search for Psychic Power: ESP and Parapsychology Revisited*. Prometheus Books.

Marks, David. (2000). *The Psychology of the Psychic.* Prometheus Books.

Park, Robert L. (2000). *Voodoo Science: The Road from Foolishness to Fraud*. Oxford U. Press.

Radin, Dean. (1997). *The Conscious Universe - The Scientific Truth of Psychic Phenomena*. HarperCollins.

Randi, James. (1982). *Flim-Flam! Psychics, ESP, Unicorns, and Other Delusions.* Prometheus Books.

Randi, James (1983a). "The Project Alpha Experiment: Part one: The First Two Years," *Skeptical Inquirer*. Summer.

Randi, James (1983b). "The Project Alpha Experiment: Part two: Beyond the Laboratory," *Skeptical Inquirer*. Fall.

Randi, James. (1997). *An Encyclopedia of Claims, Frauds, and Hoaxes of the Occult and Supernatural.* St. Martin's Press.

Williams, William. (2000). *Encyclopedia of Pseudoscience*. Facts On File, Inc.