Parapsychological Research: A Tutorial Review and Critical Appraisal

RAY HYMAN

Invited Paper

Beginning in the 1850s, some eminent scientists such as Robert Hare, Alfred Russel Wallace, and Sir William Crookes investigated the claims of spiritualist mediums and believed that they had demonstrated scientifically the existence of psychic phenomena. Critics, without examining the evidence, dismissed the claims out of hand and charged the offending scientists with gross incompetence or with fraud. Encouraged by the work of these early psychical researchers, a group of scholars founded the Society for Psychical Research in London in 1882. In spite of this beginning, psychical research remained an amateur and uncoordinated set of activities until the publication of Rhine's Extra-Sensory Perception in 1934. The card-guessing experiments featured in Rhine's book became the model for experimental parapsychology for the next 40 years. Since the 1970s Rhine's paradigm has been replaced by a number of research programs such as remote viewing, the Ganzfield experiment, and psychokinetic investigations using Random Event Generators. The present paper examines examples of what were considered, in their time, the best examples of scientific evidence for paranormal phenomena. Each generation of parapsychologists has set aside the work of earlier generations and offered up as sufficient scientific evidence the best work of its own day. As a result, parapsychology lacks not only lawful and replicable phenomena, but also a tradition of cumulative evidence. Two systematic evaluations of the best contemporary research programs in parapsychology revealed that the experiments departed from the minimal standards of adequate randomization of targets, appropriate use of statistical inference, and controls against sensory leakage. The historical survey in this paper suggests that the same themes and inadequacies that haunted the very earliest investigations still characterize contemporary parapsychological research. Both proponents and critics throughout the 130 years of the controversy over psychical research, have deviated greatly from those standards of fair-play and rationality that we would like to believe characterizes the best scientific arguments. Some encouraging signs for progress towards resolving some of the issues raised by the controversy have recently appeared. The criticism of the parapsychological claims is becoming more informed and constructive. Many younger parapsychologists have been working for higher standards within their field. The best lines of systematic research in parapsychology are not of sufficient quality to be put before the scrutiny of the rest of the scientific community. However, with the recent increase in constructive criticism and with the growing awareness within the parapsychological community that it needs to specify minimal standards and set its own house in order, there is hope that in the near future either the parapsychologists will fail to find evidence for psi or will be ready to challenge the scientific community with the sort of evidence that it cannot ignore.

Manuscript received January 25, 1985; revised August 21, 1985.
The author is with the Psychology Department, University of Oregon, Eugene, OR 97403, USA.

INTRODUCTION

Robert Jahn, Dean of the School of Engineering and Applied Science at Princeton University, can be taken as a representative example of what happens when an eminent and established scientist takes the time to carefully examine the evidence for paranormal phenomena. About seven years ago, an undergraduate requested him to supervise her investigation of psychic phenomena [1].

> Although I had no previous experience, professional or personal, with this subject, for a variety of pedagogical reasons I agreed, and together we mapped a tentative scholarly path, involving a literature search, visits to appropriate laboratories and professional meetings, and the design, construction, and operation of simple experiments. My initial oversight role in this project led to a degree of personal involvement with it, and that to a growing intellectual bemusement, to the extent that by the time this student graduated, I was persuaded that this was a legitimate field for a high technologist to study and that I would enjoy doing so.

As a result of his own survey of the field as well as his own initial experiments in parapsychology, Jahn concluded that [1]:

> once the illegitimate research and invalid criticism have been set aside, the remaining accumulated evidence of psychic phenomena comprises an array of experimental observations, obtained under reasonable protocols in a variety of scholarly disciplines, which compound to a philosophical dilemma. On the one hand, effects inexplicable in terms of established scientific theory, yet having numerous common characteristics, are frequently and widely observed; on the other hand, these effects have so far proven qualitatively and quantitatively irreplicable, in the strict scientific sense, and appear to be sensitive to a variety of psychological and environmental factors that are difficult to specify, let alone control.

Jahn, like many of his predecessors who took a serious look at the evidence for the paranormal, finds the phenomena to be erratic, evasive, and ephemeral. Indeed, he admits that when judged according to strict scientific standards, the evidence for the actual existence of the phenomena is not "fully persuasive." But he is intrigued. Like his predecessors, he is optimistic that with the right application of technology and scientific ingenuity the phenomena can be captured and made lawful.

This is one of a number of justifiable reactions one can have as a result of fairly examining the case for psychical research. Jahn is willing to risk his time and reputation on the possibility that careful and diligent investigation will bring some lawfulness to this unruly area of inquiry. Jahn's research into anomalous phenomena began over seven years ago, but it will be several more years before we know whether it has managed to progress much beyond previous attempts to bring scientific order into the field.

During the 130 year history of psychical research many other scholars and scientists initiated investigations of psychic phenomena with equally high hopes of taming the phenomena. One was the philosopher Henry Sidgwick who was the first president of the Society of Psychical Research founded in 1882. According to William James, Sidgwick and his colleagues "hoped that if the material were treated rigorously and, as far as possible, experimentally, objective truth would be elicited, and the subject rescued from sentimentalism on the one side and dogmatizing ignorance on the other. Like all founders, Sidgwick hoped for a certain promptitude of result; and I heard him say, the year before his death, that if anyone had told him at the outset that after twenty years he would be in the same identical state of doubt and balance that he started with, he would have deemed the prophecy incredible. It appeared impossible that the amount of handling evidence should bring so little finality of decision" [2].

James, who made this observation in his last article on psychical research in 1909, continued as follows [2]:

> My own experience has been similar to Sidgwick's. For twenty-five years I have been in touch with the literature of psychical research, and have had acquaintance with numerous "researchers." I have also spent a good many hours (though far fewer than I ought to have spent) in witnessing (or trying to witness) phenomena. Yet I am theoretically no "further" than I was at the beginning; and I confess that at times I have been tempted to believe that the Creator has eternally intended this department of nature to remain *baffling*, to prompt our curiosities and hopes and suspicions all in equal measure, so that, although ghosts and clairvoyances, and raps and messages from spirits, are always seeming to exist and can never be fully explained away, they also can never be susceptible of full corroboration.
>
> The peculiarity of the case is just that there are so many sources of possible deception in most of the observations that the whole lot of them *may* be worthless, and yet that in comparatively few cases can aught more fatal than this vague general possibility of error be pleaded against the record. Science, meanwhile needs something more than bare possibilities to build upon; so your genuinely scientific inquirer—I don't mean your ignoramous "scientist"—has to remain unsatisfied.

Some 67 years after James' final word on the matter, the philosopher Antony Flew summed up his 25 years of interest in parapsychology with remarkably similar sentiments [3]:

> My long-out-of-print first book was entitled, perhaps too rashly, *A New Approach to Psychical Research*. . . . When I reviewed the evidential situation at that time it seemed to me that there was too much evidence for one to dismiss. Honesty required some sort of continuing interest, even if a distant interest. On the other hand, it seemed to me then that there was no such thing as a reliably repeatable phenomenon in the area of parapsychology and that there was really almost nothing positive that could be pointed to with assurance. The really definite and decisive pieces of work seemed to be uniformly negative in their outcome.
>
> It is most depressing to have to say that the general situation a quarter of a century later still seems to me to be very much the same. An enormous amount of further work has been done. Perhaps more has been done in this latest period than in the whole previous history of the subject. Nevertheless, there is still no reliably repeatable phenomenon, no particular solid-rock positive cases. And yet there still is clearly too much there for us to dismiss the whole business.

Sidgwick was assessing the first 50 years of psychical research. James was evaluating the same period with another ten years or so added. Flew based his assessment on an additional 67 years of inquiry. Yet, all three agree that they could detect no progress. In each case, after a quarter of a century of personal involvement, the investigator found the evidence for the paranormal just as inconclusive as it had been at the beginning. James openly concedes that *all* the claimed phenomena might be the result of self-deception or fraud. Yet he, and the other two philosophers, cannot quite shake the conviction that, despite all this inconclusiveness, "there might be something there."

Over this same span of history, the critics have consistently insisted that "there is nothing there." All the alleged phenomena of telepathy, clairvoyance, psychokinesis, levitation, spirit materialization, and premonitions can be accounted for in terms of fraud, self-delusion, and simple gullibility. The proponents have naturally resented such dismissals of their claims. They have argued that the critics have not fairly examined the evidence. They have accused the critics of attacking the weakest evidence and of ignoring the stronger and better supported evidence in favor of the paranormal.

Unfortunately, as any reading of the history of psychical research quickly reveals, the psychical researchers are correct in their appraisal of their critics. Too often, the major critics have attacked strawmen and have not dealt with the actual claims and evidence put forth by the more serious researchers. The fact that most of the criticism of the psychical research has been irrelevant and unfair, however, does not necessarily mean that the psychical researchers have a convincing case.

Indeed, the message that we get from Sidgwick, James, Flew, and Jahn is that the evidential base for psychic claims is very shaky at best. At most, these scholars, after carefully weighing all the evidence available to them, are claiming only that they cannot help feeling that, despite the inconsistencies and nonlawfulness of the data, that "there must be something there."

As will be discussed later in this paper, both the critics and the proponents subscribe to what I refer to as the False Dichotomy. When a scientist or scholar, after investigating possible psychic phenomena, concludes that the phenomena are real, the assumption is that either his conclusion is justified or he is delinquent in some serious way—being either incompetent or subject to some pathology. When the critic denies that the claim is justified, the proponent feels that his integrity or competence is being challenged. And the critic, sharing in this assumption, feels that he must show that the claimant is incompetent, gullible, or deficient in some serious way [4].

I consider this a False Dichotomy because competent and honest investigators can make serious judgmental errors when investigating new phenomena. Competence and expertise in any given field of endeavor is bounded. Cognitive

psychologists, historians of science, and sociologists of knowledge have been gathering data which demonstrate how thinking is guided by conceptual frameworks and paradigms within which the thinker operates. Successful scientific thinking, for example, is not successful because it operates according to abstract, formal rules of evidence. Rather, it succeeds because the thinker is guided by the often implicit rules and procedures inherent within the specific content and practices of the narrow field of specialization within which the problem is being pursued. These "heuristics" or guidelines for successful thinking are not foolproof and under changed circumstances they can trap the thinker into erroneous convictions. In other words, competence in a given scholarly or scientific discipline and high intelligence are no barriers to becoming trapped into asserting and defending erroneous positions.

In this paper, I agree with Sidgwick, James, Flew, and Jahn in the most general sense that "something" is indeed going on. However, I do not see any need to assume that this "something" has anything to do with the paranormal.

I think we should not lightly dismiss the fact that for 130 years some of our best scholars and scientists have seriously carried out psychical research and have become convinced that they have demonstrated the existence of a "psychic force" or a supernatural realm occupied by intelligent and superior beings. As far as I can tell, these proponents were competent scholars, sane, and highly intelligent. They made every apparent effort to employ what they believed to be objective and scientific standards in observing, recording, and reporting their findings.

Yet, as I will argue, contrary to Jahn's assessment, the total accumulation of 130 year's worth of psychical investigation has not produced any consistent evidence for paranormality that can withstand acceptable scientific scrutiny. What should be interesting for the scientific establishment is not that there is a case to be made for psychic phenomena, but rather that the majority of scientists who decided to seriously investigate *believed* that they had made such a case. How can it be that so many outstanding scientists, including several Nobel Prize winners, have convinced themselves that they have obtained solid, scientific evidence for paranormal phenomena?

If they are wrong, what has made them wrong? Does this suggest weaknesses or limitations of scientific method and training? And if these investigators have not actually encountered psychic phenomena, what is it that they have discovered?

I am not sure that I can provide satisfactory answers to these questions. But I believe that it will help to look at some selected cases in which investigators believed that they had obtained adequate scientific evidence for the reality of psychic phenomena. I will start at the beginning by describing the sort of evidence that convinced the first scientists who took psychical claims seriously. Even some contemporary parapsychologists believe these early scientists may have been wrong, but their cases are still worth examining because in them we will find many of the same issues and problems that characterize contemporary parapsychological research. These early psychic investigators tested spiritualistic mediums who were noted for their ability to produce powerful psychic phenomena such as levitations, materializations, and other physical feats.

Psychical research became transformed into what is now called parapsychology when the focus shifted, after the first half century of investigation, to the study of extrasensory perception and psychokinesis in ordinary individuals by means of standardized testing materials and procedures. I will examine what was, at the time, considered to be the most rigorous and successful application of this form of parapsychological research—the now notorious investigations by Soal on Shackleton and Mrs. Stewart. Again, the purpose is not to beat a dead horse but to abstract out principles and issues that still haunt contemporary parapsychology.

The card-guessing experiments begun by Rhine in the 1930s established the paradigm which dominated parapsychology for the next 40 years. New technology and interest in altered states resulted in departures from Rhine's paradigm beginning about 1970. Experiments with Random Event Generators, Remote Viewing, and the Ganzfeld technique have been the strongest contenders for providing parapsychology with its long-sought-for repeatable experiment. I will argue that a fair and objective assessment of this latest work strongly suggests that, like its predecessors, it still does not stand up to critical scrutiny.

Scientists and Psychics

The first major scientist to test experimentally a psychic claim was Michael Faraday in 1853. As will be described in more detail in the next section, Faraday concluded that the phenomena he had investigated, table-turning, had a normal explanation. Robert Hare, a major American chemist, at first agreed with Faraday's conclusion. But, then, after personal investigations of his own, changed his mind, and openly supported the claims of spiritualistic mediums. A decade later, Alfred Russel Wallace, the cofounder with Darwin of the theory of evolution by natural selection, and Sir William Crookes, the discoverer of thallium, astounded their scientific colleagues by openly endorsing paranormal claims. Wallace and Crookes, as had Hare, believed that their own inquiries had established scientific proof to support their paranormal claims.

Hare, Wallace, and Crookes were the first of a continual succession of eminent scientists who have endorsed paranormal claims as a result of their experimental tests of alleged psychics. These scientists have established a tradition which has played a major role in the development of psychical research. The first half-century of psychical research consisted mainly of testing paranormal claims within this tradition. Beginning in the 1930s a second approach, experimental investigations according to standard protocols and using unselected subjects, became the dominant approach under the name of parapsychology. Today parapsychology includes both approaches.

In the first half of the present paper, I will focus on the first approach. The research of Sir William Crookes will be used as an example of this approach. In the second half of the paper, I will deal with the second approach. Again, I will use the research of a single investigator to bring out the more general issues and problems with the field of parapsychology. In both parts of the paper I will also briefly mention other investigators and lines of research which also bring out the same themes illustrated by the more detailed examples. Finally, I will briefly look at the contem-

porary situation in parapsychology to argue that the concerns and difficulties that haunted the earlier investigations still persist.

Table-Turning and Psychical Research

Modern spiritualism began when unaccountable raps were heard in the presence of two teen-age girls, Margaret and Kate Fox, in 1848. By using a code, the girls' mother was able to converse with the raps and concluded that they originated from the spirit of a peddler who had been murdered in the very house in which the Fox family then lived. Word of this miraculous communication spread quickly and soon a variety of means for communicating with the unseen spirits via "the spiritual telegraph" were developed in the United States and then spread to Europe. The individuals through whom the spirits produced their phenomena and communicated with mortals were called mediums. The mediums, at first, displayed phenomena such as rapping sounds, movements of tables and objects, playing of musical instruments by unseen agencies, and the occurrence of strange lights in the dark. Later, more elaborate phenomena were produced such as the levitation of objects or the medium; the disappearance or appearance of objects; the materialization of hands, faces, or even of complete spirit forms; spirit paintings and photographs; and written communications from the spirit world [5], [6].

By the early 1850s, table-turning (also called table-tilting or table-rapping) had become the rage both in the United States and in Europe. A group of individuals, usually called "sitters," would arrange themselves around a table with their hands resting flat upon the table-top. After an extended period of waiting a rap would be heard or the table would tilt up on one leg. Sometimes the table would sway and begin moving about the room, dragging the sitters along. On some occasions, sitters would claim that the table actually levitated off the floor under the conditions in which all hands were above the table. Reports even circulated that sometimes the table levitated when no hands were touching it. Table-turning was especially popular because it could occur with or without the presence of an acknowledged medium. Any group of individuals could get together and attempt to produce the phenomenon in the privacy of their own living room.

Table-turning plays an important role in the history of psychical research because it was what first attracted the attention of serious scientists to alleged paranormal phenomena [6]. The phenomenon had become so widespread in England by the summer of 1853 that several scientists decided to look into it. Although the prevailing explanation for the table's movements favored the agency of spirits, other explanations at the time were electricity, magnetism, "attraction," Reichenbach's Odyllic Force, and the rotation of the earth. Electricity, which in the public mind was then considered to be an occult and mystical force, was especially popular. Indeed, many spiritualists probably thought that the spirits operated by electricity.

In June 1853, a committee of four medical men held seances to investigate table-turning. They found that the table did not move at all when the sitters' attention was diverted and they had not formed common expectations about how the table should move. In another condition they found that the table would not move if half the sitters expected it to move to the right and the other half expected it to move to the left. "But when expectation was allowed free play, and especially if the direction of the probable movement was indicated beforehand, the table began to rotate after a few minutes, although no one of the sitters was conscious of exercising any effort at all. The conclusion formed was that the motion was due to muscular action, mostly exercised unconsciously" [6]. Other investigators came to similar conclusions.

But, by far, the most publicized and influential investigation was that by England's most renowned scientist, the physicist Michael Faraday. Faraday obtained subjects who were "very honorable" and who were also "successful table-movers" [7]. Faraday found that he could obtain movements of the table in a given direction with just one subject sitting at his table in the laboratory. His first tests were designed to eliminated as explanations well-known forces such as magnetism and electricity. He demonstrated that substances such as sand-paper, millboard, glue, glass, moist clay, tinfoil, cardboard, vulcanized rubber, and wood did not interfere with the table-turning. He could find no traces of electrical or magnetic effects. "No form of experiment or mode of observation that I could devise gave me the slightest indication of any peculiar force. No attractions, or repulsion,... nor anything which could be referred to other than the mere mechanical pressure exerted inadvertently by the turner."

Although Faraday suspected that the sitter was unconsciously pushing the table in the desired direction, the sitter adamantly insisted that he was not the agency but, instead, was pulled in the expected direction by some force within the table. Faraday created some ingenious arrangements to see if the sitter's claim was true. He placed four or five pieces of slippery cardboard, one over the other, on the table top. The pieces were attached to one another by little pellets of a soft cement. The lowest piece was attached to a piece of sandpaper which rested on the table top. The edges of the sheets overlapped slightly, and on the under surface, Faraday drew a pencil line to indicate the position. The table-turner then placed his hands upon the upper card and waited for the table to move in the previously agreed upon direction (to the left). Faraday then examined the packet. It was easy to see by displacement of the parts of the line, that the hand had moved further than the table, and that the latter had lagged behind;—that the hand, in fact had pushed the upper card to the left and that the under cards and the table had followed and been dragged by it" [7].

In another arrangement, Faraday fixed an indicator to two boards on the table top such that if the sitter was pulled by the table the indicator would slope to the right, but if the sitter pushed the table, the indicator would slope to the left. The table moved as before as long as the sitter could not see the indicator. But as soon as the sitter was able to watch the indicator, which gave him immediate feedback when his hands pushed in the expected direction, all movements of the table ceased. "But the most valuable effect of this test-apparatus... is the corrective power it possesses over the mind of the table-turner. As soon as the index is placed before the most earnest, and they perceive—as in my presence they have always done—that it tells truly whether they are pressing downwards only or obliquely; then all effects of table-turning cease, even though the parties persevere, earnestly desiring motion, till they become weary and worn out. No prompting or checking of

the hands is needed—*the power is gone*; and this only because the parties are made conscious of what they are really doing mechanically, and so are unable unwittingly to deceive themselves" [7].

Faraday's investigation convinced several scientists that table-turning was the result of self-deception resulting from unconscious motor movements guided by expectation. His report is even credited with dampening the enthusiasm, for a few years, for spiritualism in England [6]. But several spiritualists and table-turners were not convinced by Faraday's arguments. And this brings up another issue that invariably accompanies the controversy over paranormal claims. Whenever a skeptic demonstrates how an alleged psychic phenomenon can be duplicated by mundane means, the claimant usually responds, "It's not the same thing!"

To many spiritualists and those who had witnessed table-turning, Faraday's explanation appeared hopelessly inadequate. Professional mediums, for example, while sitting at the table could provide meaningful answers by means of table-rapping to questions that sitters put to their assumed spirit communicators. In addition, the table often moved in a variety of ways which seemingly could not be explained by simple muscular pressure applied by the sitters. For example, the table often levitated above the floor with all the sitters' hands resting on the top surface. And some reports claimed that the table moved and levitated when no human was in contact with it.

Faraday's explanation dealt with only one important cause of the table-turning. He did not attempt to account for the various ways in which the table could be moved and levitated by trickery. Nor did he deal with the problem of the notorious unreliability of eyewitness testimony. Nor did he and his fellow skeptics realize that an abstract, even if correct, explanation of table-turning was impotent when matched against the personal and powerfully emotional experience of a sitter who has been converted during an actual table-turning session. These same limitations on any attempt to "explain away" an alleged paranormal event by a mundane account continue to provide loopholes whereby the proponent can maintain the reality of a paranormal claim.

Two striking illustrations of the power of the experience that "it is not the same thing," can be found in the conversions to spiritualism of the next two major scientists to investigate psychic phenomena. Both Robert Hare and Alfred Russel Wallace were familiar with Faraday's research and explanation when they first investigated spiritualistic phenomena by means of table-turning. And both were immediated convinced that their personal experiences could not be accounted for by Faraday's theory. In these instances, the forewarning, rather than serving to forearm, actually disarmed. And this, too, is a recurring theme in the history of psychical research.

Sir William Crookes

Faraday, the first major scientist to seriously investigate spiritualistic phenomena, concluded that self-deception was sufficient to explain what he observed. As a result, he remained skeptical and critical of all further claims of paranormal phenomena. Faraday's scientific colleagues were obviously grateful for his investigation and conclusions. But within the next two decades three other major scientists also investigated paranormal claims and concluded, contrary to Faraday, that they had witnessed truly paranormal phenomena.

Robert Hare, the eminent American chemist, began his inquiry into spiritualistic phenomena in 1853 immediately after Faraday's investigation. Alfred Russel Wallace, the cofounder with Darwin of the theory of evolution by natural selection, initiated his investigations in 1865. And Sir William Crookes, the discoverer of thallium, began his investigations in 1869. All three had already achieved reputations as outstanding scientists before they surprised their scientific colleagues with their assertions of having witnessed psychic phenomena. Their colleagues were disturbed and puzzled by such assertions from obviously competent scientists. Their reactions, unfortunately, were not always rational and tended to make a confusing situation worse.

I believe it is important to try to understand how these otherwise competent scientists became convinced that they had acquired evidence sufficient to justify the belief in paranormal phenomena. The investigations of these scientists can be credited with the initiation of psychical research as a field with scientific aspirations. And many of the same issues of scientific justification of claims for the paranormal that we find in their work are still with us today.

Robert Hare was Professor Emeritus of Chemistry at the University of Pennsylvania and 72 years of age when circumstances conspired to launch him on a new career as a psychic investigator in 1853 [8]. Hare, the author of more than 150 scientific papers, had invented the oxy-hydrogen blowpipe which was the predecessor of today's welding torches [9]. According to Asimov, Hare was "one of the few strictly American products who in those days could be considered within hailing distance of the great European chemists" [10].

Both Hare and his critics took it for granted that a competent scientist could carry out observations and experiments on a variety of phenomena and, as a result, come to trustworthy and sound conclusions. Until he announced his conversion to the spiritualistic hypothesis, Hare's colleagues did not doubt his competence as an observer and experimenter. When he announced that he had not only experimentally verified paranormal phenomena, but had been communicating with the spirits of his departed relatives and also with George Washington, John Quincy Adams, Henry Clay, Benjamin Franklin, Byron, and Isaac Newton, this placed his incredulous colleagues in a quandary [8].

For half a century, the scientific world had accepted Hare's scientific papers and conclusions with respect and admiration. His scientific accomplishments were widely recognized and honored. But now this respected fellow scientist, by using apparently the same observational and experimental skills that had earned him his renown, was claiming to have demonstrated the reality of phenomena that scientists felt were just too preposterous to be true. Instead of examining Hare's arguments and evidence, his colleagues reacted emotionally and rejected his conclusions out of hand. Furthermore, they treated him as a traitor to the scientific enterprise and refused to allow him to present his case in the regular scientific forum.

From Hare's perspective this reaction was both unfair and unscientific. His arguments were being rejected without even being given a hearing. In his last few years he turned

away from his scientific colleagues and confined his social interactions entirely to his spiritualistic associates. From the perspective of the scientific establishment, Hare had suddenly gone insane or had suffered some other form of pathology. Here we see the False Dichotomy in action. And this same False Dichotomy will be found throughout the story of psychical research right up to the present.

Alfred Russel Wallace's conversion to spiritualism began in the same way that Hare's did—sitting at an animated table during a seance. Wallace's experience, just as Hare's did, convinced him that Faraday's explanation of the table's antics would not do. Unlike Hare, however, Wallace was not 72 and at the end of his career. Instead he was 42 years old and in the middle of a long and productive career. It had only been seven years earlier that Wallace had independently conceived the theory of evolution by natural selection, the very same theory that Darwin had been secretly working on for many years [11]–[13].

Critics have found it easy to dismiss the psychical evidence of Hare on the basis of old age and of Wallace on the assertion that, while he was a great naturalist and observer, he was not an experimenter [11]. Neither criticism can be applied, however, to William Crookes, who was the next great scientist to investigate and endorse the reality of paranormal phenomena. Crookes was generally acknowledged, even by many who opposed his psychic beliefs, as one of the preeminent chemists and physicists of his day. Crookes—the discoverer of thallium, inventor of the radiometer, developer of the Crookes tube, pioneer investigator of radiation effects, and a contributor to photography and other fields—was elected a Fellow of the Royal Society at age 31, was later knighted, and received just about every honor available to a scientist of his time.

When Crookes began attending seances with Mrs. Marshall (the same medium who helped convert Wallace) and J. J. Morse in 1869, he was 37 years of age. He had been very upset by the death of his youngest brother and apparently believed he had received spirit communications from him through the services of these mediums. In July 1870 Crookes announced his intention to conduct a scientific inquiry into spiritualistic phenomena. He wrote, "I prefer to enter upon the inquiry with no preconceived notions whatever as to what can or cannot be, but with all my senses alert and ready to convey information to the brain; believing, as I do, that we have by no means exhausted all human knowledge or fathomed the depths of all physical forces" [15].

Although most of the scientific community assumed that Crookes was undertaking the investigation as a skeptic, his biographer wrote, "But it is certain, at all events, that when in July 1870 Crookes, at the request, it is said of a London daily paper, announced his intention of 'investigating spiritualism, so-called,' he was already much inclined towards spiritualism. What he really intended to do was to furnish, if possible, a rigid scientific proof of the objectivity and genuineness of the 'physical phenomena of spiritualism,' so as to convert the scientific world at large and open a new era of human advancement" [16].

Crookes packed almost all his research into psychical phenomena into the four-year period 1870–1874 [17]. When he failed to sway his scientific colleagues—and as a result of bitter attacks by his critics, Crookes quietly dropped this work and devoted his scientific efforts from 1875 onwards to more mainstream subjects. But he never gave up his beliefs and he never severed his ties with the field. In his final years, he began attending seances again and believed, near the end, that he had finally found proof of survival when he obtained a spirit photograph of his dead wife [15].

By today's standards, the investigations that come closest to being "scientific" were those that Crookes carried out with the celebrated medium Daniel Dunglas Home. Home is probably the most colorful and enigmatic psychic in the history of spiritualism [6], [9]. In one session, which took place at Crooke's home on May 31, 1871, Home held an accordian (which had just been purchased by Crookes for this occasion) by one end so that the end with the keys hung down towards the floor. The accordian was placed in a special cage under the table which just allowed Home's hand to be inserted to hold the accordian. Home's other hand was visible above the table. The individuals sitting on either side of Home could see his hand as well as the accordian in the wire cage. "Very soon the accordian was seen by those on each side to be moving about in a somewhat curious manner, but no sound was heard..." After putting the accordian down, Home picked it up again. This time several notes were heard. Crookes' assistant crawled under the table and said that he saw the accordian expanding and contracting, but Home's hand was quite still [15].

At the same session Crookes reported an experiment that he regarded as even "more striking, if possible, than the one with the accordian." A mahogany board, 3 ft long, with one end resting on a table and other end supported by a spring balance, was in a horizontal position. Home, while "sitting in a low easy-chair" placed the tips of his fingers lightly on the extreme end of the board which was resting on the table. "Almost immediately the pointer of the balance was seen to descend. After a few seconds it rose again. This movement was repeated several times, as if by successive waves of the Psychic Force. The end of the board was observed to oscillate slowly up and down during the experiment" [15].

To see if were possible to produce an effect on the spring balance by ordinary pressure, Crookes stood on the table and pressed one foot on the end of the board where Home had placed his fingers. By using the entire weight of his body (140 lb), Crookes was able to get the index to register at most 2 lb. Home had apparently achieved a maximum displacement of 6 lb.

Because of such results Crookes concluded that, "These experiments appear conclusively to establish the existence of a new force, in some unknown manner connected with the human organisation which for convenience may be called the Psychic Force" [15]. The skeptics were not convinced. They raised a variety of objections to the experiment measuring the movement of the board. Crookes thought some of the criticisms were unfair and irrelevant. But others he felt were reasonable and could be answered.

He repeated the experiment with additional controls. To avoid direct contact with the board, he altered the apparatus slightly in a manner that had previously been used by Robert Hare in some of his experiments. A bowl of water was placed on the end of the board not supported by the spring scale. Inside the bowl of water was lowered a "hemi-

spherical copper vessel perforated with several holes at the bottom." The copper vessel was suspended from a large iron stand which was separate from the rest of the apparatus. Home placed his fingers lightly in the water in the copper bowl. Presumably, this prevented him from having direct contact with the board. Yet, under these conditions Home managed to cause the other end of the board to sway up and down.

Finally, Home was removed a few feet away from the apparatus and his hands and legs were held. Even under these conditions, Crookes was able to record movements of the board, although the displacement was less the farther Home was from the apparatus. In further answer to critics, Crookes describes similar experiments carried out successfully by other researchers including Robert Hare. Crookes also got similar results using a lady who was not a professional medium in place of Home.

This series of experiments is by far the most impressive, from a scientific viewpoint, of any that Crookes conducted. Indeed, so far as I can tell, although these were among the very first serious attempts by a scientist to test a psychic, they have not been exceeded in degree of documentation and experimental sophistication during the subsequent 114 years. This is despite the fact that following Crookes' example, eminent scientists during almost every decade since Crookes' experiments have conducted tests of famous psychics.

The comments in the preceding paragraph should not be taken as an endorsement of Crookes' results. His experiments on the "Psychic Force" are superior *relative* to what has been reported by other scientists, including contemporary ones, in their tests of psychic superstars. On an absolute scale of judgment the experiments still leave much to be desired. A major problem is documentation. Crookes omits many details which, from today's perspective at least, seem important in assessing what might have taken place.

Responding to the accusation that his witnesses were not reliable, Crookes wrote, "Accustomed as I am to have my word believed without witnesses, this is an argument which I cannot condescend to answer. All who know me and read my articles will, I hope, take it for granted that the *facts* I lay before them are correct, and that the experiments were honestly performed, with the single object of eliciting *the truth*" [15].

Here Crookes raises an important issue. When he reported finding a green line in a spectrum where one had never been reported, and followed this up with various analyses and controls to support the assertion that he must have discovered a new element (thallium), his scientific colleagues did not insist that he import skeptical witnesses, nor did they question his observations. The reported observation was made by using standard apparatus and recording procedures. The necessary controls and possibilities of error in such a context were well-known to workers in the field and it could be safely assumed that any trained chemist in this situation would behave according to both implicit and explicit rules.

But Crookes and his critics seriously err when they assume that similar confidence and trust can be placed in observations made in a field outside the investigator's training and one in which no standardization exists for instrumentation, making observations, instituting controls, recording the data, and reporting the results. The difficulties are compounded further when the observations are made, not of inanimate and reasonably passive materials, but of events involving humans who have a capacity to anticipate the experimenter's objectives and alter their behavior accordingly.

I recently discovered that Podmore, back in 1902, anticipated most of my reservations about Crookes' experiment on the movements of the balance [6]:

> The experiment as it stands, even without the modifications introduced later by Mr. Crookes in deference to his scientific critics, seems, indeed, conclusive against the possibility of Home's affecting the balance by any pressure on his end of the board. But, tested by the canons laid down by Mr. Crookes himself at the outset of his investigations, we shall find the conditions of the experiment defective in one important particular. Mr. Crookes had shown that it is the province of scientific investigation not merely to ascertain the reality of the alleged movements and measure their extent, but to establish their occurrence under conditions which render fraud impossible. In the passage quoted on page 183 it is implicitly recognised that such conditions are to be secured by eliminating the necessity for continuous observation on the part of the investigator. The proof of the thing done should depend upon something else than the mere observation of the experimenters, however skilled.
> Now in the experiment quoted these conditions were not fulfilled. On the contrary, we are expressly told that all present guarded Home's feet and hands. It is pertinent to point out that a duty for which the whole company were collectively responsible may well at times have been intermitted. Moreover, Dr. Huggins and Mr. Crookes had to watch the balance also, and Mr. Crookes had to take notes. Again, the experiment described was not the first of the kind; it occurred in the middle of a long series. It is indeed stated that Home was not familiar with the apparatus employed. But as similar apparatus had been employed, probably at previous trials by Mr. Crookes himself, certainly by earlier investigators—amongst them Dr. Hare, with whose published writings on Spiritualism we cannot assume Home was unacquainted—the statement carries little weight. Further, a point of capital importance, there had apparently been many previous trials with various modifications of the apparatus and many failures; in Mr. Crookes' own words, "the experiments I have tried have been very numerous, but owing to our imperfect knowledge of the conditions which favour or oppose the manifestations of this force, to the apparently capricious manner in which it is exerted, and to the fact that Mr. Home himself is subject to unaccountable ebbs and flows of the force, it has but seldom happened that a result obtained on one occasion could be subsequently confirmed and tested with apparatus specially contrived for the purpose."
> The real significance of this statement is that Home—a practised conjurer, as we are entitled to assume—was in a position to dictate the conditions of the experiment. By the simple device of doing nothing when the conditions were unfavourable he could ensure that the light (gas in the present instance) was such and so placed, the apparatus so contrived, and the sitters so disposed, as to suit his purpose, and that in the actual experiment the attention of the investigators would necessarily be concentrated on the wrong points. Under such conditions, as ordinary experience shows, and as the experiments described in the last chapter have abundantly demonstrated, five untrained observers are no match for one clever conjurer.

Podmore is referring, in the last sentence, to the dramatic experiments on eye-witness testimony conducted by S. J. Davey [18]. Davey had been converted to a belief in spiritualistic phenomena by the slate-writing demonstrations of the medium Henry Slade. Subsequently, Davey

accidently discovered that Slade had employed trickery to produce some of the phenomena. Davey practiced until he felt he could accomplish all of Slade's feats by trickery and misdirection. He then conducted his well-rehearsed seance for several groups of sitters, including many who had witnessed and testified to the reality of spiritualistic phenomena. Immediately after each seance, Davey had the sitters write out in detail all that they could remember having happened during his seance. The findings were striking and very disturbing to believers. None of the sitters had suspected Davey of using trickery. Sitters consistently omitted crucial details, added others, changed the order of events, and otherwise supplied reports which would make it impossible for any reader to account for what was described by normal means.

Podmore has much more to say about this experiment. His reference to "untrained" observers is not meant to question Crookes' scientific competence. "But his previous training did not necessarily render him better qualified to deal with problems differing widely from those presented in the laboratory. To put it bluntly, if Home was a conjurer, Mr. Crookes was probably in no better position for detecting the sleight-of-hand than any other man his equal in intelligence and native acuteness of sense. Possibly even in a worse position; for it may be argued that his previous training would prepare the way for Home's efforts to concentrate attention on the mechanical apparatus, and thus divert it from the seemingly irrelevant movements by which it may be conjectured the conjurer's end was attained."

Finally, Podmore points out ways in which the report is incomplete. He then speculates about one possible way Home might have tricked Crookes. He describes a scenario in which Home could have employed a thread which he attached to the apparatus, probably the hook of the scale. Some further points could be mentioned such as the fact that Crooke's unpublished notes suggest that the experiment was much more informal and involved many more distractions than the published version indicates [15].

Crookes held many seances not only with Home but with almost every major spiritualistic medium who was in England during the years 1869 through 1875. He reported having observed a variety of phenomena which he argued could not have been produced by normal means: movement of heavy bodies with contact but without mechanical exertion; raps and other sounds; the alteration of weights of bodies; movements of heavy substances at a distance from the medium; the rising of tables and chairs off the ground, without contact of any person; the levitation of human beings; the appearance of hands, either self-luminous or visible by ordinary light; direct writing; and phantom forms and faces [18]. His documentation for such phenomena, however, falls far short of what he has supplied us for the movements of the balance.

As was the case with Hare and Wallace, Crookes was bitterly attacked for his views. The eminent physiologist, William Carpenter, lead the opposition. Carpenter openly questioned Crookes' competence as a scientist, wrongly stated that Crookes' election to the Royal Society had been questionable, and made several other unwarranted insults [16], [17]. Like Wallace, Crookes tried to get his scientific colleagues and critics to witness his experiments with Home and other psychics. But none of them accepted his invitations.

Difficulties in Testing Alleged Psychics

Hare, Wallace, and Crookes were the first of many eminent scientists who have investigated and endorsed psychics. Their work inspired many later scientists to also take time away from their regular scientific activities to investigate the paranormal claims of mediums or self-professed psychics [4], [19]–[29]. Yet, I suspect that many parapsychologists will object to using the work of these psychic investigators as part of a general evaluation and critique of parapsychology. The objection would be based on two arguments.

Today, most parapsychologists would not include the reports of Hare, Wallace, and Crookes in their case for the reality of psi (the current term to refer to extrasensory perception and psychokinesis). And, secondly, even the reports by more recent scientists on psychics do not form part of the primary database of parapsychology. Instead, today's parapsychologists want to base their argument on evidence emerging from laboratory experiments with unselected subjects and which use standardized tasks.

However, I believe there are good reasons for focussing on these early investigators:

1) At the time they were reported, these investigations were considered to be the strongest evidence for the paranormal. From 1850 to 1866 Hare's research constituted practically the entire "scientific" case upon which proponents could base their claims. From 1870 until the founding of the Society of Psychical Research in 1882, it was the work of Crookes and Wallace that proponents put forth as the best scientific justification for their paranormal claims.

2) The psychical research of these three eminent scientists served as the model for all later investigations of psychics by scientists. Although sometimes the latest technological developments are brought into the investigations, no change in approach or improvements in methodology for such investigations has occurred during the 130 years since Hare first reported his findings [23]. In terms of adequacy of documentation, for example, it is difficult to find any improvement over Crookes' reports on his experiments with Home in the subsequent accounts by such psychic investigators as Richet, Barrett, Lodge, Lombroso, Zoellner, Eisenbud, Targ, Puthoff, Hasted, and the many others.

3) The work of this early trio served as an important impetus for the subsequent founding of the Society for Psychical Research in 1882. In his presidential address to the first general meeting of The Society for Psychical Research on July 17, 1882, Henry Sidgwick went out of his way to acknowledge the importance and evidential value of the work of these pioneer researchers [30]:

> I say that important evidence has been accumulated; and here I should like to answer a criticism that I have privately heard which tends to place the work of our Society in a rather invidious aspect. It is supposed that we throw aside *en bloc* the results of previous inquiries as untrustworthy, and arrogate to ourselves a superior knowledge of scientific method or intrinsically greater trustworthiness—that we hope to be believed, whatever conclusions we may come to, by the scientific world, though previous inquirers have been uniformly distrusted. Certainly I am conscious of making no assumption of this kind. I do not presume to suppose that I could produce evidence better in quality than much that has been laid before the world by writers of indubitable scientific repute—men like Mr. Crookes, Mr. Wallace, and the late Professor de Morgan. But it is clear that from what I

> have defined as the aim of the Society, however good some of its evidence may be in quality, we require a great deal more of it. I do not dispute, it is not now time to dispute with any individual who holds that reasonable persons, who have looked carefully into the evidence that has been so far obtained, ought to be convinced by that evidence; but the educated world, including many who have given much time and thought to this subject, are not yet convinced, and therefore we want more evidence.

Sidgwick makes it clear that he and the other founders of the Society for Psychical Research consider the findings of Wallace and Crookes as scientifically sound. Sidgwick has no doubt that Wallace's and Crookes' reports *should* convince reasonable members of the scientific community. But he pragmatically makes the distinction between what *should* and what *will* convince the critics. "What I mean by *sufficient evidence* is evidence that will convince the scientific world, and for that we obviously require a good deal more than we have so far obtained" [30]. In other words, Sidgwick does not aspire to improve the quality of the preceding scientific investigators. Rather he wants to acquire more of the same quality.

4) The investigations of these original psychical researchers bring out many of the same issues of evidence, testimony, and proof that still characterize current controversies in parapsychology. Unfortunately, not much in the way of further clarification or resolution of these issues has occurred since their efforts first stimulated the debate. I have already mentioned some of these issues in my discussions of the individual cases.

Many of the issues involve the problem of competency. To what extent, for example, does competency in one branch of inquiry transfer, if at all, to a different branch? Can a scientist, no matter how competent and well-intentioned, initiate an inquiry into a previously unstructured and unstandardized area and single-handedly produce results which bear the same scientific status as the results he has produced in his original area of expertise? Elsewhere, I have given by reasons for answering this question in the negative [23].

One important issue is perhaps worth bringing up at this point. The scientists who have defended the trustworthiness of their psychical research have typically insisted that the observations and evidence of their reports of psychic happenings do not differ in quality from that which characterizes their more orthodox investigations.

Yet, at the same time, these same investigators acknowledge an important difference between their inquiries into physics and biology and their investigations of psychics. Hare, Wallace, and Crookes, as well as the later psychical researchers insisted that the psychics being tested must be treated with proper respect and concern for their feelings. If the investigator is overly skeptical or otherwise betrays distrust of the alleged psychic this could adversely affect the paranormal performance. Thus these scientists try to convey the impression that they conduct their tests using every precaution against fraud and deception, but at the same time making sure not to take any step or include any condition that meets with the disapproval of the alleged psychic. Skeptics such as myself, who have both experience in conducting experiments with humans and have been trained in conjuring, believe this is an impossible task. The twin goals of preventing trickery on the part of the alleged psychic and of ensuring that this same person will be satisfied with all the experimental arrangements are mutually incompatible.

But scientists who have testified to the paranormal powers of their subjects confidently insist they have simultaneously achieved both goals. A contemporary version of this theme has been eloquently put forth by a group of scientists, including two of England's outstanding physicists, in describing their experiments on the psychokinetic powers of Uri Geller [31]:

> We have come to realize that in certain ways the traditional ideal of the completely impersonal approach of the natural sciences to experimentation will not be adequate in this domain. Rather, there is a personal aspect that has to be taken into account in a way that is somewhat similar to that needed in the disciplines of psychology and medicine. This does not mean, of course, that is not possible to establish facts on which we can count securely. Rather, it means that we have to be sensitive and observant, to discover what is a right approach, which will properly allow for the subjective element and yet permit us to draw reliable inferences. One of the first things that reveals itself as one observes is that psychokinetic phenomena cannot in general be produced unless *all* who participate are in a relaxed state. A feeling of tension, fear, or hostility on the part of any of those present generally communicates itself to the whole group. The entire process goes most easily when all those present actively want things to work well. In addition, matters seem to be greatly facilitated when the experimental arrangement is aesthetically or imaginatively appealing to the person with apparent psychokinetic powers.
>
> We have found also that it is generally difficult to produce a predetermined set of phenomena. Although this may sometimes be done, what happens is often surprising and unexpected. We have observed that the attempt to concentrate strongly in order to obtain a desired result (e.g., the bending of a piece of metal) tends to interfere with the relaxed state of mind needed to produce such phenomena.... Indeed, we have sometimes found it useful at this stage to talk of, or think about, something not closely related to what is happening, so as to decrease the tendency to excessive conscious concentration on the intended aim of the experiment....
>
> In the study of psychokinetic phenomena, such conditions are much more important than in the natural sciences, because the person who produces these phenomena is not an instrument or a machine. Any attempt to treat him as such will almost certainly lead to failure. Rather, he must be considered to be one of the group, actively cooperating in the experiment, and not a "subject" whose behavior is to be observed "from the outside" in as cold and impersonal manner as possible....
>
> In such research an attitude of mutual trust and confidence is needed; we should not treat the person with psychokinetic powers as an "object" to be observed with suspicion. Instead, as indicated earlier, we have to look on him as one who is working with us. Consider how difficult it would be to do a physical experiment if each person were constantly watching his colleagues to be sure that they did not trick him. How, then, are we to avoid the possibility of being tricked? It should be possible to design experimental arrangements that are beyond any reasonable possibility of trickery, and that magicians will generally acknowledge to be so. In the first stages of our work we did, in fact, present Mr. Geller with several such arrangements, but these proved to be aesthetically unappealing to him. From our early failures, we learned that Mr. Geller worked best when presented with many possible objects, all together on a metal surface; at least one of these objects might appeal to him sufficiently to stimulate his energies....
>
> Nevertheless, we realize that conditions such as we have described in this paper are just those in which a conjuring trick may easily be carried out. We understand also that we are not conjuring experts, so if there should be an intention to deceive, we may be as readily fooled as any person.

> Moreover, there has been a great deal of public criticism, in which the possibility of such tricks has been strongly suggested. For this reason it has often been proposed that a skilled magician should be present to help to see that there will be no possibility of deception. It is in the nature of the case, however, that no such assurance can actually be given. For a skilled magician is able to exploit each new situation as it arises in a different and generally unpredictable way... . In principle, we would welcome help of this kind in decreasing the possibility of deception. It has been our observation, however, that magicians are often hostile to the whole purpose of this sort of investigation, so they tend to bring about an atmosphere of tension in which little or nothing can be done. Indeed, even if some magicians who were found who were not disposed in this way, it does not follow that their testimony will convince those who are hostile, since the latter can always suppose that new tricks were involved, beyond the capacity of those particular magicians to see through them. Because of all of this, it seems unlikely that significant progress towards clearing up this particular question could be made by actually having magicians present at the sessions, though we have found it useful to have their help in a consultative capacity... . We recognize that there is a genuine difficulty in obtaining an adequate answer to criticisms concerning the possibility of tricks, and that a certain healthy skepticism or doubt on the part of the reader may be appropriate at this point... . However, we believe that our approach can adequately meet this situation.

These investigators close this discussion of the difficulties of carrying out such research with an optimistic prognosis, "We feel that if similar sessions continue to be held, instances of this kind might accumulate, and there will be no room for reasonable doubt that some new process is involved here, which cannot be accounted for, or explained, in terms of the laws of physics at present known. Indeed, we already feel that we have very nearly reached this point." These hopeful words were written in 1975. Neither they nor other scientists have yet managed to present scientific evidence that Uri Geller or his many imitators can bend metal paranormally. Although at least one major physicist continues his investigations of paranormal metal bending [20], a decade of research on Uri Geller by scientists who adhered to the advice of treating the metal-bender as a respected colleague and catering to his aesthetic sensibilities has only succeeded to demonstrate that Geller can bend metal under conditions which allow him to do it by cheating [21].

Hare, Wallace, and Crookes, as well as subsequent psychic researchers, insisted they had guarded against the possibility of trickery while, at the same time, acknowledging the necessity to treat their subjects in the special way described by Hasted *et al.* Unfortunately, as Hasted *et al.* concede, this special treatment increases the difficulties of preventing deception. But, like their predecessors in psychical research, they express confidence that their scientific skills can overcome the difficulty. In fact, the suggested procedure gives the alleged psychic veto power over any arrangement that impedes trickery and also supplies a ready excuse for not producing phenomena when the dangers of detection suddenly seem too high. The conditions which the scientists report as ideal for the production of psychical phenomena are just those that are also ideal for the production of the same phenomena by trickery.

5) As already discussed, Hare, Wallace, and Crookes were bitterly attacked by their skeptical scientific colleagues. And the same sorts of attacks and defenses have characterized subsequent cases. Both critics and defenders still implicitly subscribe to the same False Dichotomy. And both the critics and the defenders, in different ways, do not emerge as rational, objective, scientific or otherwise admirable in their exchanges. Worse, no lessons from the past seem to have either been learned or carried over to the current controversies. If the critical exchanges had been more constructive and rational at the time of Hare, Wallace, and Crookes, today we might be closer to understanding what was really going on to make such eminent scientists put forth such seemingly outrageous claims.

Hare, Wallace, and Crookes had no success in inducing their critics to come and examine the evidence for themselves. It is possible that if Huxley and Carpenter had accepted Wallace's invitation to attend at least six seances, no phenomena would have taken place. On the other hand, it would be useful to have the accounts of such skeptical observers before us if, say, Miss Nichol did produce the flowers in their presence. And it certainly would have helped if Carpenter and Stokes had accepted Crookes' invitation to watch his experiments with Home and the balance.

The Creery Sisters

For its first 30 years, psychical research consisted of individual and uncoordinated investigations by scholars or scientists such as Hare, Wallace, and Crookes. During this period some feeble and unsuccessful attempts were made to form research societies and coordinate the research [32]. The first successful attempt to institutionalize psychical research was the founding of the Society for Psychical Research in London in 1882. Four of the principal leaders of this society—the philosopher Henry Sidgwick, the physicist William Barrett, the literary scholar Edmund Gurney, and the classicist Frederic Myers—had been encouraged, in addition to their own investigations of telepathy and mediums, by the research of such scientists as Wallace and Crookes. The founders of the Society clearly believed that they possessed solid scientific evidence for the reality of thought-transference. At the first general meeting of the Society in London on July 17, 1882, Henry Sidgwick ended his presidential address with the following words [30]:

> We must 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 a blindness or forgetfulness incompatible with any intellectual condition except absolute idiocy. I am glad to say that this result, in my opinion, has been satisfactorily attained in the investigation of thought-reading. Professor Barrett will now bring before you a report which I hope will be only the first of a long series of similar reports which may have reached the same point of conclusiveness.

Before looking at the experimental results whose "conclusiveness" Sidgwick believes is beyond reasonable doubt, I would like to call the reader's attention to the use of the False Dichotomy in Sidgwick's strategy. The goal is to report evidence that is so compelling that the critic either has to admit that psychic phenomena have been demonstrated or that the investigator is deliberately lying, afflicted with a pathological condition, or incredibly incompetent. Sidgwick does not allow for the possibility that an investigator could be competent, honest, sane, and intelligent, and still wrongly report what he believes to be "conclusive"

evidence for the paranormal. Unfortunately, as seen in the cases of Hare, Wallace, and Crookes and as typifies succeeding the cases, the critics, in responding to paranormal claims, have implicitly accepted the False Dichotomy. When confronted with paranormal claims by otherwise competent investigators, many critics have taken the bait and have tried to discredit the offending investigator by questioning his competence, insinuating fraud, or suggesting pathology.

The "conclusive" evidence with which Sidgwick wanted to confront the objector came from a series of experiments on through-transference conducted by his colleagues William Barrett, Edmund Gurney, and Frederic Myers [33]. The investigators introduced this series as follows [33]:

> In the correspondence we have received there were two cases which seemed, upon inquiry, to be free from any *prima facie* objections, and apparently indicative of true thought-reading. One of these cases is given in the Appendix . . . but as we cannot from personal observation testify to the conditions under which the trials were made, we simply leave it aside. The other case was that of a family in Derbyshire, with whom we have had the opportunity of frequent and prolonged trials.
> Our informant was Mr. Creery, a clergyman of unblemished character, and whose integrity indeed has, it so happens, been exceptionally tested. He has a family of five girls, ranging now between the ages of ten and seventeen, all thoroughly healthy, as free as possible from morbid or hysterical symptoms, and in manner perfectly simple and childlike. The father stated that any one of these children (except the youngest), as well as a young servant-girl who had lived with the family for two years, was frequently able to designate correctly, without contact or sign, a card or other object fixed on in the child's absence. During the year which has elapsed since we first heard of this family, seven visits, mostly of several days' duration, have been paid to the town where they live, by ourselves and several scientific friends, and on these occasions daily experiments have been made.

The preceding quotation was taken from the "First Report on Thought-Reading" which was read at the first meeting of the Society. Several more experiments were conducted with the Creery sisters and the results included in the second and third reports [34], [35]. Notice the emphasis placed upon Reverend Creery's "unblemished character" and integrity. Within the Victorian society of Sidgwick and his colleagues this emphasis on character had a special significance. According to Nicol, many flaws in the investigative reports of the Society were due to "a *double standard* of evidence."

> The Society's double standard of evidence arose in the following way. The Society's leaders were members of the middle and upper middle strata of society. When faced with the problem of estimating the value of evidence, they divided the world into two classes: (a) Members of their own class (Ladies and Gentlemen in the Victorian sense) whom they tended to treat trustingly; (b) Members of the lower classes, whom for brevity we may call the Peasants: them they treated with suspicion [36].

The experiments with the Creery sisters were all variants of the popular Victorian pastime known as the "willing game" [37].

> The game admits of many variations, but is usually played somewhat as follows. One of the party, generally a lady, leaves the room, and the rest determine on something which she is able to do on her return—as to take a flower from some specified vase, or to strike some specified note on the piano. She is then recalled, and one or more of the 'willers' place their hands lightly on her shoulders. Sometimes nothing happens, sometimes she strays vaguely about; sometimes she moves to the right part of the room and does the thing, or something like the thing, which she has been willed to do. Nothing could at first sight look less like a promising starting-point for a new branch of scientific inquiry.

Barrett, Gurney, and Myers go to great lengths to assure their readers that they are aware of the many nonparanormal ways in which information from the senders can be communicated to the percipient. Subtle unconscious pushes by the "willer," for example, can guide the percipient to the correct place. And there is always the possibility of secret codes being employed [33], [37]. Nevertheless, they relate incidents from their own experience with the game which they believe cannot be handled by such obvious explanations.

In their typical experimental procedure, one child would be selected to leave the room. When she was out of the room, the remaining participants would select a playing card or write down a number or name. "On re-entering she stood—sometimes turned by us with her face to the wall, oftener with her eyes directed towards the ground, and usually close to us and remote from her family—for a period of silence varying from a few seconds to a minute, till she called out to us some number, card, or whatever it might be" [33]. Before leaving the room, the child was always informed of the general category, such as playing cards, from which the target item was to be chosen.

The authors obviously felt that their knowledge of the various ways that inadvertent and deliberate signaling of the percipient could occur somehow made them immune from such errors. As an added precaution, however, they conducted several trials either in which members of the family were absent or in which only the experimenters knew the chosen object (unfortunately they do not distinguish among trials on which only the experimenters were informed of the target but the family was present and trials on which only the experimenters were present). The investigators claim that keeping the family uninformed did not appreciably lower the proportion of above-chance correct guesses.

The results were quite striking. Looking only at the results of those trials on which members of the Committee alone knew the card or number selected, the investigators summarize their findings as follows [35]:

> 260 Experiments made with playing cards; the first responses gave 1 quite right in 9 trials; whereas the responses, if pure chance, would be 1 quite right in 52 trials. 79 Experiments made with numbers of two figures; the first responses gave 1 quite right in 9 trials; whereas the responses, if pure chance, would be 1 quite right in 90 trials.

The experimenters also summarize the results of the much larger number of trials in which the family members were not excluded. Two points are worth noting about the results reported above. By ordinary statistical criteria the odds against such an outcome being due just to chance are enormous. But the calculation of such odds assumes, that in the absence of telepathy, we know the expected value and distribution of hits. The way experimenters can ensure the appropriate conditions for the application of the statistical tests is to include careful procedures for randomizing the

targets on each trial such that each target has an equal chance of being selected and that the selected object on a given trial is independent of the selection on the next. But nowhere in the three reports do we find any mention of how the playing card or number was chosen on each trial. We do not know if the deck was shuffled even once, let alone between trials. The number selection is even more disturbing because if, as seems to be the case, a committee member simply thought of any two digit number that came to mind, we know that some numbers are much more likely than others. And the same few numbers that are favored by the sender are likely to be those that come to the mind of the percipient. These most probable numbers, known as "mental habits" in the older literature, are called "population stereotypes" by Marks and Kammann [25].

The second peculiarity, which was noted by Coover, is that the proportion of successful hits in these experiments seems to be independent of the chance probability [38]. Thus the hit rate is 1 out 9 trials regardless of whether cards or numbers are being guessed. To Coover this suggest the use of a code rather than the imperfect transmission of psychic signals.

As already indicated, the founders of the Society for Psychical Research believed that, with the experimental results on the Creery sisters, they had finally succeeded in scientifically establishing telepathy as a valid phenomenon. As just one example of the importance attached to these experiments, Gurney's statement in the Society's first major monograph, *Phantasms of the Living* [39] can be cited:

> I have dwelt at some length on our series of trials with the members of the Creery family, as it is to those trials that we owe our own conviction of the possibility of genuine thought-transference between persons in a normal state.

Despite this confidence in the conclusiveness of the Creery experiments, critics quickly pointed out perceived flaws [38], [40], [41]. It was charged that the authors grossly underestimated the extent to which sophisticated coding could transpire between the girls in the experimental situation. The critics also suggested that the experimenters were naive in assuming that they could prevent inadvertent cueing just by being aware of the possibility.

Concerning the trials in which only the investigators knew the chosen object, the critics complained about inadequate documentation. The experimenters never state how the card or object was chosen; whether the members of the family were present during the selection (even though they were presumably kept ignorant of the choice); whose deck of playing cards was used; and so forth.

As can be seen, even on this brief account, we encounter a number of the issues that characterized earlier psychical research. The investigators assume that to be forewarned is to be forearmed. For example, they devote six pages of their first report to a discussion of the various types of errors, which if not excluded, could invalidate their research [33]. The purpose is to assure the reader that because they are keenly aware of the possibilities of such errors they could not have occurred. As previously mentioned, one way the investigators tried to preclude giving the girl any involuntary muscular cue was simply for the investigator to be consciously aware of such a possibility and consciously prevent himself from displaying such cues. Not only is such a precaution useless [42], but it was unnecessary since one could more directly prevent unwitting bodily cues by simply screening those who know the target from the percipient. This tendency to substitute *plausible* (to the investigator) reasons for discounting a possible source of error for actual experimental controls to guard against the error characterizes psychical research from its inception to the present.

A second theme is that prior experience in investigating paranormal claims automatically qualifies one as an expert who can be trusted not to make mistakes or be susceptible to trickery in future situations. This theme is closely related to the False Dichotomy issue.

The report on the Creery sisters also illustrates another recurring theme in psychical research—the *Patchwork Quilt Fallacy*. As Giere points out, the "patchwork quilt fallacy" gets its name because, "The hypothesis, initial conditions, and auxiliary assumptions are pieced together in such a way that they logically imply the known facts" [43]. Telepathy or psi always seems to be just that mysterious phenomenon that produced all the peculiar patterns that we happened to observe in our data. On some days the Creery sisters performed no better than chance. This variability among days became, in the minds of the investigators, a property of the phenomenon [35]:

> It may be noted that the power of these children, collectively or separately, gradually diminished during these months, so that at the end of 1882 they could not do, under the easiest conditions, what they could do under the most stringent in 1881. This gradual decline of power seemed quite independent of the tests applied, and resembled the disappearance of a transitory pathological condition, being the very opposite of what might have been expected from a growing proficiency in code-communication.

The fact that alleged psychics inevitably seem to lose their powers under continued investigation has become known as the "decline effect," which can occur in a variety of patterns and guises. Gurney and his colleagues propose the decline as additional support for the genuineness of the telepathy because it is not what might be expected if the girls were becoming more proficient in using a code. The cynic, of course, views this decline in the just the opposite way. Presumably the investigators are also becoming more proficient in knowing what to look for, especially in the face of continuing criticism, and, as a result, they have made it more difficult for the girls to get away with their tricks.

As it turns out the investigators later caught the girls cheating. The girls, at least on this occasion, had used a simple code. This brings up an additional theme in psychical research which we might, for short, label the *Problem of the Dirty Test Tube*. Gurney revealed the deception in a brief note which appeared in the *Proceedings of the Society for Psychical Research* in 1888 [44]. Hall thinks it is very significant that Gurney's fellow investigators did not sign this revelation [41].

In the note, Gurney reminds his readers "that the earliest experiments in Thought-transference described in the Society's *Proceedings* were made with some sisters of the name of Creery. The important experiments were, of course, those in which the 'agency' was confined to one or more of the investigating Committee.... But though stress was never laid on any trials where a chance of collusion was afforded by one or more of the sisters sharing in the

'agency,' nevertheless some results contained under such conditions were included in the records. It is necessary, therefore, to state that in a series of experiments with cards, recently made at Cambridge, two of the sisters, acting as 'agent' and 'percipient,' were detected in the use of a code of signals; and a third has confessed to a certain amount of signalling in the earlier series to which reference has been made" [44]. Gurney then describes both the visual and auditory codes used by the girls. He continues as follows [44]:

> The use of the visual code was very gratuitous on the part of the sisters, since it had been explained to them that we did not attach any scientific value to the experiments in which they acted as agent and percipient in sight of each other, the possibility of success under these conditions having been abundantly proved. The object of our experiments at Cambridge on this occasion was, if possible, to strengthen the evidence for Thought-transference (1) when no members of the family were aware of the thing to be guessed, and (2) when the sister acting as agent was in a different room from the one acting as percipient. The experiments in which the codes were used were intended merely as amusement and encouragement with a view to increase the chance of success in the more difficult ones—which were all complete failures. The account which was given as to the earlier experiments, conducted under similar conditions, is that signals were very rarely used; and not on specially successful occasions, but on occasions of failure, when it was feared that visitors would be disappointed. But of course the recent detection must throw discredit on the results of all previous trials in which one or more of the sisters shared in the agency. How far the proved willingness to deceive can be held to affect the experiments on which we relied, where collision was excluded, must of course depend on the degree of stringency of the precautions taken against trickery of other sorts —as to which every reader will form his own opinion.

This manner of treating the discovery of cheating illustrates a number of interwoven themes. The finding of a "dirty test tube" ordinarily implies that all the results of the experiment are brought into question. Gurney argues that only those results clearly attached to the "dirty test tube" should be discarded. Since the girls could not have used their code, in his judgment, in those trials in which only investigators knew the chosen object, those trials still retain their evidential value. Related to this is what the early psychical researchers called the problem of "mixed mediumship." Psychics and mediums are under constant pressure to produce results, yet they have little direct control over their fickle powers. Therefore, in order not to disappoint their followers or from fear of losing the attention that goes with mediumship, they learn to supplement their real powers with tricks to simulate the phenomena. Still another variant of this exploits the apparent fact that many mediums and psychics are apparently in a trance or altered state when performing. In such a state they are highly suggestible and behave in ways expected of them. If skeptics are among the onlookers, they will sometimes cheat because this is what is expected of them. The onus for the consequent cheating is by this means placed upon the skeptic rather than the cheater.

The dirty test tube problem has been with psychical research from its beginning and, as we will see, is still very much a part of the contemporary scene. The medium Eusapia Palladino's long career was noteworthy for the number of times she was caught cheating. She readily acknowledged that she would cheat if the investigators gave her the opportunity. Despite this record of cheating, many psychical researchers, including some of today's leaders in the field, have no doubt that on many other occasions she displayed true paranormal powers [19]. On the contemporary scene, parapsychologists are willing to admit that the controversial metal-bender Uri Geller often cheats, but that, on occasion, he exhibits real paranormal powers [45]. And parapsychologists blamed me, rather than Geller, for the fact that Geller cheated in my presence because, as they put it, I did not impose sufficiently stringent conditions to prevent him from cheating [22].

Despite this attempt to save some of the evidence from the Creery experiments, the leaders of the Society for Psychical Research quietly removed the experiment from their evidential database. But Sir William Barrett refused to go along with this demoting of the experiment. According to Gauld, this incident sparked dissension between Barrett and the other founders [32].

> Barrett had been the first to experiment with these girls, and they were his special proteges.... Barrett would never agree that the later and crude cheating invalidated all the earlier results; he considered that his 1876 experiments, together with his experiments with the Creerys had established his claim to be the discoverer of thought-transference, and he remained bitter towards the Sidgwicks for the rest of his life.

Not only did Barrett continue to defend the evidential value of the Creery experiments, but so did later parapsychologists. In his classic monograph of 1934 on *Extra-Sensory Perception*, J. B. Rhine included this experiment as among the most evidential of the early research. "On the whole the early experiments in E.S.P. were admirably conducted...as one would expect from the array of highly impressive names connected with them. The experiments with the Creery sisters, for instance, were conducted by Professors William Barrett, Henry Sidgwick and Balfour Stewart, by Mrs. Henry Sidgwick, Frederic Myers, Edmund Gurney and Frank Podmore.... In all this work the results were sufficiently striking to leave no doubt as to the exclusion of the hypothesis of chance" [46].

Despite these attempts to salvage something from the Creery experiments, I believe it is fair to say that today the experiments are not part of the case that parapsychologists would make in support of psi. Indeed, my perusal of several contemporary books and histories of parapsychology indicates that the experiments are rarely, if ever, mentioned.

The same fate befell the very next major experiment on telepathy conducted by the same investigators. In their "Second Report on Thought-Transference," Gurney and his colleagues describe the first of their experimental findings in which two young men, Smith and Blackburn, were apparently able to communicate telepathically under conditions that prevented normal communication. If anything, the investigators placed even more reliance upon these later experiments than in those with the Creery sisters.

As was the case with the Creery sisters, Smith and Blackburn soon lost their powers. Smith was then hired by the Society to assist in the conduct of several successful telepathic experiments. In 1908, Blackburn, thinking that Smith was dead, publicly confessed as to how he and Smith had tricked the investigators during the experiments. Smith, who was very much alive and still employed by the Society, denied the charges. In the ensuing debate, the Society's

leaders defended Smith. Good accounts of this amazing incident can be found in [38] and [41]. Today, the Smith–Blackburn experiments are no longer considered part of the parapsychological case for psi.

J. B. RHINE

The founding of the Society for Psychical Research in 1882 was an attempt to organize and professionalize psychical research. Other societies, such as the American Society for Psychical Research quickly followed. Journals and proceedings were published and international congresses were held. Despite these steps towards institutionalization, psychical research continued for the next 50 years to be an uncoordinated activity of amateurs. No agreed upon program or central body of concepts characterized the field.

During this period, psychic researchers disagreed among themselves on issues involving subject matter, methodology, and theory. On one side were those, perhaps the majority, who supported the spiritist hypothesis that psychic phenomena reflected the activity of departed spirits or superintelligent beings. Opposed to these were psychic researchers like Nobel Laureate Charles Richet who defended the position that the phenomena could be explained in terms of a "psychic force" without assuming survival or spirits [47].

Another division was between those who felt that psychical research should confine itself to mental phenomena such as telepathy, premonitions, and clairvoyance. Opposed to these were those who felt that the physical phenomena such as levitation, materialization, poltergeist events, and psychokinesis should be the focus of inquiry. The majority of psychical researchers believed in telepathy but were dubious about clairvoyance. But a strong minority, lead by Richet, believed that clairvoyance not only existed but was the basic phenomenon underlying telepathy.

Possibly the most divisive issue of all was the question of what sort of a research program was appropriate for psychical investigation. A small, but vocal minority wanted psychical research to become a rigorous experimental science. A larger group felt that the natural-historical method was more appropriate because so many of the important phenomena were spontaneous and not observable in the laboratory. Opposed to both these groups were members of the societies who felt that the quantification and rigor of the natural sciences were irrelevant to the study of psychical phenomena.

The event that is credited with providing psychical research with a common focus and a coherent research program was the publication in 1934 of J. B. Rhine's monograph *Extra-Sensory Perception* [46]. Mauskopf and McVaugh [47] provide an excellent survey of the period from 1915 to 1940, which they treat as the period when psychical research made the transition from a pre-paradigmatic to a paradigmatic research program.

Rhine pulled together the various strands already existing in psychical research and coordinated them into a coherent program. He also coined the terms "parapsychology" to refer to the new experimental science which descended from psychical research and "extra-sensory perception" to refer to the basic phenomenon which was to be studied: In agreement with Richet, and in disagreement with the British parapsychologists, Rhine viewed clairvoyance as on the same footing with telepathy. Later, precognition was also put under the rubric of extra-sensory perception (ESP). ESP became defined as "Knowledge of or response to an external event or influence not apprehended through known sensory channels" [48]. This included telepathy, clairvoyance, precognition, and retrocognition. The psychic phenomena not involving reception of information were included under the term "psychokinesis" (PK) which is defined as "The influence of mind on external objects or processes without the mediation of known physical energies or forces" [48]. Today both ESP and PK are included under the more general term "psi" which is "A general term to identify a person's extrasensorimoter communication with the environment" [48].

Rhine's 1934 mongraph deals only with clairvoyance and telepathy. In 1934 he also began research programs on precognition and psychokinesis. Apparently, he was reluctant to publicize these latter programs too soon for fear of making parapsychology too controversial and unacceptable to mainstream science [48]. He waited until 1938 before he published anything on precognition and until 1943 for the first reports on his PK results.

The major innovation introduced by Rhine was the use of the five target designs: circle, cross, wavy lines, square, and star. These patterns were printed on cards and the standard ESP deck consisted of 5 cards of each symbol for a total of 25 cards. Rhine also introduced standard procedures for using these target materials. The two most common were the Basic Technique and the Down Through Technique. In the Basic Technique (B.T.), the deck is shuffled and placed face down, the percipient guesses the value of the top card; this is then removed and laid aside and the percipient guesses the value of the second card; the second card is then removed and laid on top of the first and the percipient now guesses the third card; etc. This procedure is continued until all 25 cards have been used. At the end of such a "run," a check is made to see how many guesses were hits. If the procedure was supposed to test telepathy then an agent would look at each card at the time the percipient was trying to guess its symbol. If clairvoyance was being tested, no one would look at each card as it was placed aside. The Down Through Technique (D.T.) tested clairvoyance by having the percipient guess the symbols from top to bottom before any of them were removed for checking against the call. The D.T. technique is considered to be superior methodologically in that it better protects against inadvertent sensory cues from the backs of the cards.

Extra-Sensory Perception attracted the attention of both the psychical researchers and the skeptics for two reasons. Rhine's database consisted of 91 174 separate trials or guesses over a three-year period using a number of nonprofessional individuals as percipients. More important was the unprecedented level of success which he reported. Of the 85 724 guesses recorded using the five-symbol ESP decks, 24 364 were "hits." This was 7219 more hits than the 17 145 that would be expected just by chance. The odds against this being just an accident are calculated as being practically infinite. His subjects averaged 7.1 hits per run of 25 as against the chance expectation of 5. Although this is only 2 extra hits per 25, such consistency over this huge number of trials and different subjects had no precedent in the prior history of psychical research.

Rhine's best subject, Hubert Pearce averaged 8 hits per run over a total of 17 250 guesses. As Rhine notes [46]:

> Most people are more impressed by a spectacular series of successive hits than by lower but cumulative scoring. Pearce's scoring 25 straight hits under clairvoyant conditions, in my presence, and Zirkle's 26-straight hits in pure telepathy with my assistant, Miss Ownbey, are the best instances of these. Other subjects have approached these. Linzmayer scored 21 in 25 clairvoyance, in my presence; Miss Ownbey herself, unwitnessed, scored 23, pure clairvoyance. Miss Turner's score of 19 in distance P.T. [pure telepathy] work stands out because of the 250 miles between her and the agent. Miss Bailey scored 19 in P.T. in the same room with the agent, as did also Cooper. The odds against getting one series of 25 straight hits by mere chance would be 5^{35} which is nearly 300 quadrillions—just one score of 25! A small part of our 90 000 trials.

Rhine's work provided the model for most parapsychological work from 1934 to around 1970. Using card-guessing with the five ESP symbols, an astonishing variety of questions about ESP were investigated [48]. Because of its huge database, its claims to statistical and experimental sophistication, and its unprecedented rate of success Rhine's research gained the attention of scientific and popular audiences [47]. At first scientists were at a loss about how to react. Many scientists, as a result of reading Rhine's work, were encouraged to try to replicate the results. A few got encouraging results, but most failed.

The first attacks by the critics were aimed at Rhine's statistical procedures. As it turned out, some of Rhine's statistical procedures were technically incorrect, but, for the most part, his results could not be explained away as due to inappropriate statistical procedures. The critics turned out to be wide off the mark in many of their accusations. On the whole, however, the statistical debate led to constructive developments and improved clarification about the proper use of statistical procedures in such experiments [47].

Having essentially lost the statistical battle, the critics then turned to Rhine's experimental controls. Here, he was much more vulnerable. And, ironically, it was the British psychical research community that had anticipated the critics and which provided the sharpest critiques of Rhine's methods [47]. The British parapsychologists were astonished both by Rhine's apparent ease in finding successful percipients as well as his claims that clairvoyance worked as well as telepathy. With only a few exceptions, they had found only evidence for telepathy. And their experience had convinced them that telepathic powers were very rare. While they welcomed Rhine's contribution, they were quick to point out many of its defects, especially Rhine's inadequate description of his procedures and the seeming casualness of his experiments.

During the 1930s, nevertheless, Rhine's work as reported in *Extra-Sensory Perception*, was hailed by parapsychologists as the best scientific case for ESP ever put before the world. Today, as I understand it, most parapsychologists, although they acknowledge its seminal influence on the development of the field, dismiss much of Rhine's earlier work as nonevidential because of its loose controls, poorly made target materials, and inadequate documentation.

S. G. Soal

Rhine's strongest critic among the British parapsychologists was the mathematician S. G. Soal. Just prior to the appearance of Rhine's monograph, Soal had conducted a huge series of card-guessing experiments with only chance results. But the experiments for which Soal became most renowned began as a direct response to Rhine's monograph.

After five years of heroic research, Soal was sure that he had succeeded only in demonstrating the laws of chance. A colleague, however, persuaded him to check for a certain trend in his data. And this resulted in a new series of experiments that for almost 25 years were hailed as the most convincing and fraud-proof demonstration of ESP ever achieved. Because the experiment and results seemed so impressive, some critics, in a way reminiscent of Carpenter's attacks upon Wallace and Crookes and within the spirit of Sidgwick's False Dichotomy, openly accused Soal of fraud on no other basis than that his results were too good. Other critics attacked him on grounds that were irrelevant. As it turns out the critics were right, but for the wrong reasons!

As soon as Soal heard about Rhine's successful American research, he began an ambitious program to replicate Rhine's findings in England. Soal started late in 1934 and continued his experiments for five years. At the end he had accumulated 128 350 guesses for 160 percipients. This is almost 30 percent more guesses than Rhine had accumulated for his 1934 monograph. Soal was sure that he had removed all the flaws and weaknesses that had characterized Rhine's work. Unfortunately, Soal found that this enormous effort yielded "little evidence of a direct kind that the persons tested, whether considered as individuals or in the mass, possessed any faculty for either clairvoyance or telepathy" (quoted in [49]).

Soal reported these results to a stunned parapsychological world in 1940. At the same time another British parapsychologist, Whately Carington, reported the results of telepathy experiments which seemed to show a "displacement effect." Instead of achieving hits on the target, his subjects seemed to achieve above chance matches when their guesses were matched with either the immediately preceding or the next target in the series. Carington asked Soal to check his data to see whether he, too, might find such a displacement effect [49].

Soal was reluctant to do so. He told Goldney that he thought Carington's request was preposterous and he wasn't going to waste his time going through his huge batch of records. But Carington persisted and Soal finally agreed. Soal found, among the records of his 160 percipients, two who seemed to show Carington's displacement effect. Although this finding was published, presumably Soal realized that such a post hoc finding had to be replicated [49].

Fortunately, one of his two percipients, Basil Shackleton, was available for testing during the years 1941 through 1943. With the collaboration of K. M. Goldney, 40 sittings which yielded a total of 11 378 guesses were obtained with Shackleton during this difficult period when England was at war. As had been the case with the original testing, Shackleton's guesses were at chance level when compared with the actual target, but when compared with the symbol coming up immediately after the target (precognitive hitting), Shackleton's guesses yielded 2890 successes as compared with the 2308 expected by chance. The odds against this being a chance occurrence were calculated to be more than 10^{35} to 1 [50].

In 1945 Soal was able to begin experimenting on the second percipient who had displayed the displacement effect in the original data, Mrs. Gloria Stewart. He was able to accumulate a total of 37 100 guesses during 130 separate sittings. Unlike Shackleton's or her own previous perfor-

mance, her hitting this time was on the actual target rather than on the immediately preceding or following trial. She managed to achieve 9410 hits which were 1990 more hits than would be expected by chance. The odds against such a result were calculated as 10^{70} to 1 [50].

Soal's stated objective was to make these experiments completely error-free and fraudproof. The basic procedure, which was varied slightly on occasion, was as follows. The percipient—Basil Shackleton or Gloria Stewart—sat in one room monitored by one of the experimenters (EP). In an adjoining room, the sender or agent sat at table opposite the second experimenter (EA). The door between the rooms was slightly open so that the percipient could hear EA's call as to when to make his or her guess. The percipient, of course, could see neither the agent or EA. A screen, with a small aperture separated the agent and EA. For each block of 50 trials EA had before him a list of randomized numbers which determined the target for each trial. Each number could range from 1 to 5. If the target number for the first trial was, say, 3, EA would hold up a card with the number 3 on it so that it could be seen by the agent through the aperture. The agent had lying before him in a row, five cards. Each card had a different drawing of an animal on it: elephant, giraffe, lion, pelican, and zebra. Before each block of trials, the agent shuffled the order of the picture cards. If EA held up a card with 3 on it, the agent would turn up the third card and concentrate upon the animal depicted on it. The percipient would then try to guess which animal was being "sent" and write his guess for that trial in the corresponding place on the response sheet. After every block of 50 trials, the agent reshuffled the target cards so that, for that block, only the agent knew which animal corresponded with which number.

In addition to this rather elaborate arrangement, independent observers were invited to attend many of the sittings. Several professors and a member of parliament were among the observers. On some blocks of trials, unknown to the percipient, the agent did not look at the symbols. This was a test for clairvoyance. Other variations were introduced from time to time. The experiments with Gloria Stewart, while following the same pattern, were admittedly not as carefully controlled. Special precautions were also introduced to ensure that the prepared target sequences could not be known to agent or percipient in advance. And careful safeguards were introduced during the recording of the results and the matching of the targets against the guesses. Duplicates of all records were made and posted immediately after each session to a well-known academic.

Never before had so many safeguards been introduced into an ESP experiment. With so many individuals involved, and with prominent observers freely observing, any form of either unwitting cueing or deliberate trickery would seem to be just about impossible. If fraud of any sort were to be suspected, it would seemingly require, under the stated conditions, the active collusion of several prominent individuals. Beyond these safeguards, Soal randomized his targets, instituted sophisticated checks for randomness, and used the most appropriate statistical procedures. Despite these elaborate precautions, the two subjects managed to consistently score above chance over a number of years.

Soal's findings were hailed as definitive by the parapsychological community and were so good that the rest of the scientific community, including the skeptics, could not ignore them. Here was one of Rhine's severest critics, a man who had spent many years meticulously conducting enormous card-guessing experiments with only chance results, a man who was by profession a mathematician, and an experimenter who had seemingly taken every known precaution to guard against every loophole and possibility of error, who suddenly demonstrated highly successful telepathic and precognitive results over sustained periods of time with two percipients.

Whately Carington, the parapsychologist who convinced Soal to re-examine his seemingly unsuccessful results, wrote (as quoted in [51]):

> Mr. Soal is a most remarkable man, for whose work I have the highest possible admiration. Possessed of a more than Jobian patience, and a conscientiousness, thoroughness which I can only describe as almost pathological, he worked in various branches of the subject for many years with nothing but a succession of null results to show for it... Hoping to repeat Rhine's experiments in England, he tested 160 persons, collecting 128 350 Zener card guesses singlehanded, and using the most elaborate precautions against every possible source of error... If I had to choose one single investigation on which to pin my whole faith in the reality of paranormal phenomena, or with which to convince a hardened skeptic (if this be not a contradiction in terms), I should unhesitatingly choose this series of experiments, which is the most cast-iron piece of work I know, as well as having yielded the most remarkable results.

Similar sentiments were expressed by virtually every parapsychologist who commented on this work. As just one illustration, R. A. McConnell [52] phrased it as follows:

> As a report to scientists this is the most important book on parapsychology since the 1940 publication of *Extra-Sensory Perception After Sixty Years*. If scientists will read it carefully, the 'ESP controversy' will be ended.

G. R. Price's Critique

Ironically, some critical scientists did read it carefully, but, contrary to McConnell's prognosis, the controversy did not end. Indeed, one of the first major reviews in a scientific journal raised the controversy to new heights. Although the Shackleton experiments had originally been reported by Soal and Goldney in the *Proceedings of the Society for Psychical Research* in 1943, the scientific world did not become aware of those experiments until they were reported along with the later experiments with Gloria Stewart in the 1954 book *Modern Experiments in Telepathy* by Soal and Bateman [50].

What fueled the controversy was an unprecedented review article, nine pages in length, appearing in *Science*, the prestigious journal of the American Association for the Advancement of Science. On August 26, 1955 George R. Price's article on "Science and the Supernatural" was the only feature article for that issue. Price, who as far as I can tell had never before written on parapsychology, was described as being a research associate in the Department of Medicine at the University of Minnesota.

Price began his controversial article by stating that, "Believers in psychic phenomena—such as telepathy, clairvoyance, precognition, and psychokinesis—appear to have won a decisive victory and virtually silenced opposition" [53]. Price writes that such a victory has seemed close in the past, but always critics have managed to find flaws. But

Price sees the time at which he is writing as unique because practically no scientific papers had attacked parapsychology during the preceding 15 years [53].

> The victory is the result of an impressive amount of careful experimentation and intelligent argumentation. The best of the card-guessing experiments of Rhine and Soal show enormous odds against chance occurrence, while the possibility of sensory clues is often eliminated by placing cards and percipient in separate buildings far apart. Dozens of experimenters have obtained positive results in ESP experiments, and the mathematical procedures have been approved by leading statisticians.
>
> I suspect that most scientists who have studied the work of Rhine (especially as it is presented in *Extra-Sensory Perception After Sixty Years*,... and Soal (described in *Modern Experiments in Telepathy*),... have found it necessary to accept their findings.... Against all this evidence, almost the only defense remaining to the skeptical scientist is ignorance, ignorance concerning the work itself and concerning its implications. The typical scientist contents himself with retaining in his memory some criticism that at most applies to a small fraction of the published studies. But these findings (which challenge our very concepts of space and time) are—if valid—of enormous importance, both philosophically and practically, so they ought not to be ignored.

Price then elaborates upon a suggested scheme, using redundancy coding, which would make ESP useful, even if it is a very weak and erratic form of communication. He then presents his version of Hume's argument against miracles. He quotes Tom Paine's more succinct version of the same argument, "...Is it more probable that nature should go out of her course, or that a man should tell a lie?"

To justify using Hume's argument as his only grounds for accusing the parapsychologists of cheating, Price first tries to show that if ESP were real it would violate a number of fundamental principles underlying all the sciences. Some of these principles are that the cause must precede the effect, signals are attenuated by distance, signals are blocked by appropriate shielding, and so forth. ESP, according to Price, if it exists, violates all these principles. Then Price puts forth reasons why he considers ESP to be a principle of magic rather than merely a previously undiscovered new law of nature. "The essential characteristic of magic is that phenomena occur that can most easily be explained in terms of action by invisible intelligent beings... The essence of science is mechanism."

These lengthy considerations back up Price's solution to coping with the challenge of parapsychological claims [53]:

> My opinion concerning the findings of the parapsychologists is that many of them are dependent on clerical and statistical errors and unintentional use of sensory clues, and that all extrachance results not so explicable are dependent on deliberate fraud or mildly abnormal mental conditions.

Actually, nothing is novel or startling about Price's opinion. The same opinion, stated in just about the same words, probably is held by all skeptics. Price has carried his opinion beyond skepticism, however. The thrust of his article is that the *best* research in parapsychology as exemplified in the work of Rhine and Soal *cannot* be dismissed on the basis of "clerical and statistical error and unintentional use of sensory clues." Therefore, he concludes that the results of this otherwise exemplary research *must* be due to fraud. He does not feel that he requires any evidence of fraud. Hume's argument against miracles gives him sufficient license. Price's position, of course, no longer belongs to skepticism, but rather to dogmatism. His position seemingly is that no research, no matter how well done, can convince him of ESP.

But Price does not want to go to quite that extreme. He says that he still can be convinced *provided that* the parapsychologists can supply him with just one successful outcome from a truly fraudproof experiment. "What is needed is one completely convincing experiment—just one experiment that does not have to be accepted simply on the basis of faith in human honesty. We should require evidence of such nature that it would convince us even if we knew that the chief experimenter was a stage conjurer or a confidence man."

But does not the Soal experiment with Shackleton and Stewart meet this criterion? No, says Price, because he can imagine scenarios in which cheating could have taken place. Price then presents a number of possible ways that he feels cheating *could* have occurred in the Soal experiments [53].

> I do not claim that I know how Soal cheated if he did cheat, but if I were myself to attempt to duplicate his results, this is how I would proceed. First of all, I would seek a few collaborators, preferably people with good memories. The more collaborators I had, the easier it would be to perform the experiments, but the greater would be the risk of disclosure. Weighing these two considerations together, I'd want four confederates to imitate the Shackleton experiments. For imitating the Stewart series, I'd probably want three or four—although it is impossible to be certain, because the Stewart sittings have not been reported in much detail. In recruiting, I would appeal not to desire for fame or material gain but to the noblest motives, arguing that much good to humanity could result from a small deception designed to strengthen religious belief.

After providing a sampling of scenarios in which cheating could have occurred, all involving the collusion of three or more investigators, participants and onlookers, Price supplies some designs of what he would consider to be a satisfactory test. The key to all his designs involves a committee. "Let us somewhat arbitrarily think of a committee of 12 and design tests such that the presence of a single honest man on the 'jury' will ensure validity of the test, even if the other 11 members should cooperate in fraud either to prove or disprove occurrence of psi phenomena."

Perhaps if some enterprising group of scientists collaborated and conducted an ESP experiment with positive results according to one of Price's approved designs, the outcome might very well convince *him*. But I do not think it would, nor should it, convince the majority of skeptical scientists. Without going into all its other faults, a single experiment—no matter how elaborate or allegedly fraudproof—is simply a unique event. Scientific evidence is based on cumulative and replicable events across laboratories and investigators. The rubbish heap of scientific history contains many examples of seemingly air-tight experiments whose results have been discarded because later scientists could not replicate the results. The experiments on mitogenetic radiation would be just one example. No one has found fault with the original experiments. But since later experimenters could not replicate the results, the original experiments have been cast aside. Can anyone doubt that this would not also happen to a successful, but nonreplicable, ESP outcome from one of Price's "satisfactory tests?"

Price tells us, "that I myself believed in ESP about 15 years ago, after reading *Extra-Sensory Perception After Sixty Years*, but I changed my mind when I became acquainted with the argument presented by David Hume in his chapter 'Of miracles' in *An Enquiry Concerning Human Understanding*." So Hume supplies him with his escape hatch.

But all this seems unnecessarily dramatic. Price has fallen into a particularly stark version of the False Dichotomy. He has been forced into the very position that Henry Sidgwick wanted for the critics. The best ESP evidence is so good that either the critic must admit the reality of psi or accuse the proponents of lying and fraud. In falling into this trap, one that critics from the days of Hare and Crookes right up to the present keep falling into, Price has needlessly attributed to the Rhine and Soal results a level of evidential value which they cannot carry. At the same time, Price has implied that he is sufficiently expert in parapsychological research that he can infallibly judge when a given outcome unquestionably supports the conclusions of the experimenters. In fact, I doubt that even the parapsychologists are ready to give such power to a single experiment, even one so seemingly well-conducted as Soal's.

Price writes as if, when confronted with experimental evidence for psi, such as can be obtained by reading *Extra-Sensory Perception After Sixty Years* or *Modern Experiments in Telepathy*, he must immediately a) find ways to reject the findings on the basis of possible sensory leakage, statistical artifacts, or loose experimental controls; or b) accept the outcome as proof of psi; or c) accuse the investigators of fraud if he can imagine some scenario, no matter how complex and unlikely, under which fraud *could* have occurred. Price just does not understand either parapsychological research or scientific research in general if he truly believes these are the only alternatives open to him. Unfortunately, Price is behaving like many of the other outspoken critics of psychical research. To Price's credit, he has at least tried to make his basis for action explicit.

Both Rhine and Soal, in their responses to Price's critique, eagerly accepted Price's implicit endorsement of their experimental procedures. Soal commented that, "It is very significant and somewhat comforting to learn that Price admits that 'most of Soal's work' cannot be accounted for by any combination of statistical artifact and sensory leakage" [54]. Soal also examined in detail Price's various proposed schemes for faking the experiments [54]:

> Price goes to great length in devising variations on this theme, but they all depend on the Agent being in collusion with the chief Experimenter or with the Percipient. Now four of the Agents with whom Mrs. Stewart was highly successful were lecturers of high academic standing at Queen Mary College in the University of London. Two were senior lecturers and the other two were mathematicians who had done distinguished creative work. A fifth Agent who was brilliantly successful over a long period was a senior civil servant, in fact an assistant director of mathematical examinations in the Civil Service. Now is it plausible to suppose that I, as chief Experimenter, could persuade any of these men to enter into a stupid and pointless collusion to fake the experiments over a period of years? What had any of them to gain from such deplorable conduct? If I had gone to any of them and suggested (as Price recommends) that in a good cause a little deception would do no harm, I know quite plainly that the result would have been a first-class scandal in university circles.

Rhine found even more solace in Price's attack. "Strange though it may seem, the publication of the George Price paper . . . is, on the whole, a good event for parapsychology" [55]. For one thing, it was a way of getting a lot of instruction on parapsychology before the scientific community. Rhine also felt Price's vivid portrayal of the potential importance of ESP was valuable. He welcomed Price's effective rebuttal against the standard criticisms against ESP. And Rhine especially liked the fact that Price focussed on the point that psi was incompatible with the materialism of science [55]:

> [Price], even more than any other critical reviewer, gives indication of having felt the force of the evidence for ESP. When he turns then—albeit a bit too emotionally—and says that, according to the current concept of nature, ESP is impossible and therefore the parapsychologists must all be fakers, he at least draws the issue where it can be squarely met. The answer of the parapsychologist is: "Yes, either the present mechanistic theory of man *is* wrong—that is, fundamentally incomplete—or, of course, the parapsychologists *are* all utterly mistaken." *One* of these opponents is wrong; take it, now, from the pages of *Science*! This recognition of the issue gives point to the findings of parapsychology in a way none can easily miss.

Notice that Rhine and Price agree on some aspects of this controversy. Both Rhine and Price believe that if the claims of parapsychology are correct the foundations of science are seriously threatened. Rhine welcomes such a destruction of what he calls materialism. Price seems willing to take the most drastic measures to avoid this overthrow of what he calls the basic limiting principles. (Not all parapsychologists agree with Rhine that the acceptance of psi need be inconsistent with scientific materialism.) One issue involves what it means for contemporary science to accept the reality of psi. This concerns matters that are currently controversial among philosophers of science. And so, it is probably not fruitful to attempt to deal with them here.

Rhine and Price also agree that the standard arguments against parapsychological evidence do not hold up. According to reasonable scientific criteria, the evidence for psi is more than adequate. And so it is at this point that both Rhine and Price want to have the showdown. Price, as a defender of the materialistic faith, puts all his money on the hope that the parapsychologists have faked the data. He has no evidence to back this claim. But if he can invent possible scenarios wherebye trickery *might* have been committed in a given experiment, then he believes he can, under license from David Hume, assume that fraud must have taken place. He is not completely dogmatic about this. If the parapsychologist can come up with positive results in at least one experiment conducted under what Price considers to be fraudproof conditions, then Price has committed himself to accept the consequences.

Many issues are raised by Price's dramatic confrontational posturing. At this point, I will just mention one. Price goes beyond conventional scientific practice when he empowers a given experiment with the ability to prove the existence of psi. Once we realize that no experiment by itself definitely establishes or disproves a scientific claim, then Price's extreme remedies to save his image of science become unnecessary. No matter how well-designed and seemingly flawless a given experiment, there is always the possibility that future considerations will reveal previously unforseen loopholes and weaknesses.

Indeed, a careful analysis of the Soal experiment will reveal a variety of weaknesses. For example, in spite of the number of observers and experimenters, Soal always had control over the prepared target sequences or over the

basic recording. And both Shackleton and Stewart only produced successful results when Soal was present. On one occasion, without informing Soal, his co-investigator Mrs. Goldney conducted a sitting with Shackleton. The outcome was unsuccessful. The American parapsychologist J. G. Pratt ran a series of experiments with Mrs. Stewart without Soal's presence. No evidence for psi was found. And whereas all Rhine's results showed no difference between telepathic and clairvoyance trials, both Shackleton and Mrs. Stewart produced successful results only on telepathic trials. Furthermore, in spite of the much vaunted measures to guard against sensory leakage, the actual experimental setup, when carefully considered, offered a variety of possibilities for just such unwitting communication.

None of the foregoing considerations, in themselves, account for Soal's findings. But they make superfluous, I would argue, the hasty assumption that the findings can only be explained either by psi or some elaborate form of dishonest collusion.

The Discrediting of Soal

As it turns out, if Soal did cheat—and it now seems almost certain that he did, he almost certainly did so in ways not envisaged by either Price or Hansel. The scenarios generated by these two critics involved collusion among several of the principals. Soal apparently managed the fraud entirely on his own, or, at most, with the collusion of one other person. Furthermore, he probably used a variety of different ways to accomplish his goals.

If it had not been for a series of seemingly fortuitous events, Soal's experiment might still occupy the honored place in the parapsychologists' exhibits of evidence for psi [56]–[60]. The discrediting of Soal's data occurred through a number of revelations during the period from 1955 through 1978. Up until 1978 the accumulation of evidence suggested that something was highly suspicious about the records in the Shakleton experiments. The case was strong enough to discredit Soals' results in the judgment of some leading parapsychologists, but many others still defended Soal's findings.

The final blow to the credibility of Soal's results came in 1978 when Betty Markwick published her article "The Soal-Goldney experiments with Basil Shackleton: New evidence of data manipulation" [60]. As with the previous revelations of peculiarities in the data, Markwick's stunning findings arose out of a series of fortuitious incidents.

The story is much too complicated to relate here. Essentially, Markwick had begun a rather elaborate project to clear Soal of the accumulating charges that he had tampered with the data. Her plan involved searching the records with the aid of a computer to find subtle patterns which, if they existed, would account for the anomalies found by the critics and would vindicate Soal. Markwick did not find such patterns. Instead, she discovered previously unnoticed patterns that could be accounted for if one assumed that Soal had used a sophisticated plan for inserting "hits" into the records while he was apparently summarizing and checking the results. Reluctantly, she was forced to conclude that only the hypothesis of deliberate tampering with the data could explain her findings [60].

> Protestations to the effect that Soal, a respected scientist, would not have cheated in his own experiments—and that anyway the rigorous experimental conditions in the Shackleton series precluded fraud—seem to me to carry little weight in the face of the evidence. We can rarely fathom how conjurors achieve *their* feats, and perhaps Soal was as clever. It is futile to argue that the prison cell is escape-proof when the inmate has clearly gone.

Markwick, obviously dismayed at having discovered that Soal almost certainly faked his data, suggests two possible explanations for why he might have done so. One of her hypotheses made use of the well-known fact that Soal sometimes did automatic writing in a dissociated state. Markwick suggested the possibility that Soal may have had a split personality and that the cheating was done by his other self.

Markwick's second hypothesis involved data massage and has more universal psychological plausibility (although it is not necessarily inconsistent with her first hypothesis). She assumes that Soal's enormous accumulation of negative ESP findings were obtained legitimately. She also assumes that his *post hoc* finding of consistent displacement effects in the data of Basil Shackleton and Gloria Stewart was also legitimate [60].

> Having embarked upon the Shackleton series, one may imagine the scoring rate begins to fade (as ESP scores are wont to do after the initial flush of success). Soal, seeing the chance slipping away of gaining scientific recognition for Parapsychology, a cause in which he passionately believes, succumbs to the temptation of "rectifying" a "temporary" deficiency.

Markwick's second scenario is consistent with known patterns in which scientists have tampered with their data [61], [62]. The components appear to be: 1) the investigator believes, on the basis of previous experience, that the phenomenon under investigation is "real"; 2) for some unknown reason his current research fails to reveal the phenomenon; 3) if he reports negative results his readers might wrongly believe that the phenomenon does not exist; 4) as a result, the "truth" and assumed positive consequences of the phenomenon might be lost to humanity. Given these ingredients, it takes a very small step for the investigator to convince himself that he is helping both the truth and a good cause along by doctoring his data.

William James, with reference to his experiences in psychical research, suggested that cheating in order to convince others of the "reality" you know to be the case might be defensible. James discussed this matter in his last essay on psychical research. He referred to the policy of English investigators to consider a medium who has been caught cheating as one who always cheats. He indicated that he thought this had generally been a wise policy [2].

> But, however wise as a policy the S.P.R.'s maxim may have been, as a test of truth I believe it to be almost irrelevant. In most things human the accusation of deliberate fraud and falsehood is grossly superficial. Man's character is too sophistically mixed for the alternative of "honest or dishonest" to be a sharp one. Scientific men themselves will cheat—at public lectures—rather than let experiments obey their well-known tendency towards failure.

James gave two examples of such cheating. And then revealed the following about his own behavior [2]:

> To compare small men with great, I have myself cheated shamelessly. In the early days of the Sanders Theater at Harvard, I once had charge of a heart on the physiology of which Professor Newell Martin was giving a popular lecture. This heart, which belonged to a turtle, supported an index-straw which threw a moving shadow, greatly enlarged, upon the screen, while the heart pulsated. When certain nerves were stimulated, the lecturer said, the heart would act in

certain ways which he described. But the poor heart was too far gone and, although it stopped duly when the nerve of arrest was excited, that was the final end of its life's tether. Presiding over the performance, I was terrified at the fiasco, and found myself suddenly acting like one of those military geniuses who on the field of battle convert disaster into victory. There was no time for deliberation; so, with my forefinger under a part of the straw that cast no shadow, I found myself impulsively and automatically imitating the rhythmical movements which my colleague had prophesied the heart would undergo. I kept the experiment from failing; and not only saved my colleague (and the turtle) from humiliation that but for my presence of mind would have been their lot, but I established in the audience the true view of the subject. The lecturer was stating this; and the misconduct of one half-dead specimen of heart ought not to destroy the impression of his words. "There is no worse lie than a truth misunderstood," is a maxim which I have heard ascribed to a former venerated President of Harvard. The heart's failure would have been misunderstood by the audience and given the lie to the lecturer. It was hard enough to make them understand the subject anyhow; so that even now as I write in cool blood I am tempted to think that I acted quite correctly. I was acting for the *larger* truth, at any rate, however automatically... To this day the memory of that critical emergency has made me feel charitable towards all mediums who make phenomena come in one way when they won't come easily in another. On the principles of the S.P.R., my conduct on that one occasion ought to discredit everything I ever do, everything, for example, I may write in this article—a manifestly unjust conclusion.

I wonder if James would have approved of the way William Crookes covered up the cheating of the medium Mary Showers in behalf of "the larger truth?" Mary Showers, a young medium, conducted at least one joint seance with Florence Cook in Crookes' home. Apparently Crookes had several other sittings with Mary. Daniel Home presumably heard rumors that Crookes might be having an affair with the young Mary Showers. Crookes wrote a letter to Home explaining how the scandal had originated [63].

According to Crookes he had obtained a complete confession from Mary Showers in her own handwriting that her phenomena were wholly dependent upon trickery and the ocassional use of an accomplice. Crookes said, however, that he had undertaken not to reveal the fact that Mary was fraudulent even to her own mother, because of "the very great injury which the cause of truth would suffer if so impudent a fraud were to be publicly exposed."

The Post-Rhine Era

Rhine's card-guessing paradigm dominated experimental parapsychology from 1934 to at least the 1960s. Since the 1960s card-guessing experiments have played a minor role. Contemporary parapsychologists have deviated from Rhine's paradigm in a variety of ways. In Rhine's paradigm both the possible targets and the possible responses are severely restricted. The targets consist of five, deliberately neutral and simple, symbols. And, on each trial, the percipient is restricted to calling out the name of one of these possible five symbols. From a strictly methodological viewpoint these restrictions have several advantages. Most percipients have no strong preferences for any of the symbols; randomizing of targets is straightforward; scoring of hits and misses is unambiguous; and the statistical calculations are fairly standard.

But these same features have been blamed by contemporary investigators for the lack of impressive findings since the spectacular scoring reported by Rhine in 1934 [46]. Because the symbols are relatively meaningless and uninteresting, the repetitive guessing over many trials is boring and, according to the parapsychologists, contributes to both a lack of motivation and emotional involvement which might be needed for the effective functioning of psi.

As a result, one break with the past is the increased use of more complex and meaningful targets such as reproductions of paintings, travel slides, geographical locations, and emotionally laden photographs. In addition, instead of the forced-choice procedure of the card-guessing, most experimenters allow free-responding on the part of their percipients. Percipients are encouraged, on a given trial, to free-associate and describe, both in words and in drawings, whatever comes to mind. The use of free responses complicates enormously the problems of scoring and statistical analysis. But parapsychologists believe the added complications are a small price to pay if the newer procedures produce better psychic functioning.

Along with the free-response designs, parapsychologists have renewed their interest in the possibility that psychic functioning may be enhanced in altered states such dreaming, hypnosis, meditation, sensory-deprived state and progressive relaxation. The basic idea is that these altered states greatly reduce or block attention to external sensory information while, at the same time, increasing attention to internal mentation. Under such conditions it is hypothesized that the psi signal is easier for the percipient to detect because it has less competition from sensory inputs [64]. One survey of 87 experiments in which percipients were in an altered state found that 56 percent reported significant hitting of targets [65].

Another departure from the Rhine paradigm was stimulated by developments in electronic technology. Psi experiments employing Random Event Generators began in the 1970s. Electronic equipment could be used to generate random targets as well as automatically record the percipient's responses and keep running tallies of the hits. Although such equipment has been used to test ESP, the most widespread use has been in the study of psychokinesis. In such experiments an operator or "psychic" attempts to bias the output of a random event generator by mental means alone. In 1980, May, Humphrey, and Hubbard found reports of 214 such experiments, "74 of which show statistical evidence for an anomalous perturbation—a factor of nearly seven times chance expectation" [66].

A third major departure has been the so-called "Remote Viewing" paradigm [22], [24], [28], [67], [68]. The claims made for the ability of this procedure to consistently demonstrate ESP with a variety of percipients are perhaps the strongest ever put forth by parapsychologists [28].

Our laboratory experiments suggest to us that anyone who feels comfortable with the idea of having paranormal ability can have it.... In our experiments, we have never found anyone who could not learn to perceive scenes, including buildings, roads, and people, even those at great distances and blocked from ordinary perception.... We have, as of this writing, carried out successful remote viewing experiments with about twenty participants, almost all of whom came to us without any prior experience, and in some cases, with little interest in psychic functioning. So far, we cannot identify a single individual who has not succeeded in a remote viewing task to his own satisfaction.

In a more recent assessement of remote viewing, Targ and Harary assert, "In laboratories across this country, and

in many other nations as well, forty-six experimental series have investigated remote viewing. Twenty-three of these investigations have reported successful results and produced statistically significant data, where three would be expected" [68].

A fourth emphasis has been the study of personality correlates of the alleged psi ability [48].

In addition to the experimental programs on altered states, random event generators, remote viewing, and personality correlates contemporary parapsychologists have been actively doing research in other areas. The various chapters in the *Handbook of Parapsychology* provide a good idea of the range of topics [48]. The research on reincarnation, survival after death, paranormal photography, psychic metal bending, poltergeist phenomena, hauntings, and faith healing, while admittedly colorful, does not deserve the serious attention of scientists—at least not in its current state. I suspect that most serious parapsychologists would also not want to rest their case on such research.

Today the parapsychologists who want the scientific establishment to take their work seriously do not offer for inspection the evidence that previous generations of psychic researchers believed was sufficient—the findings of Hare, Wallace, Crookes, Gurney, Rhine, or Soal. Nor do they offer up the reports on reincaration, psychic healing, paranormal photography, spoon bending, psychic detection, and the related phenomena which so readily appeals to the media and the public. Instead, they ask us to look at the trends and patterns which they find in research programs carried out in a variety of different parapsychological laboratories.

Two aspects of this new type of claim are worth noting. One is the admission that a single investigation, no matter how seemingly rigorous and fraud-proof, cannot be acceptable as scientific evidence. The idea of a single "critical experiment" is a myth. The second, and related, aspect is that replicability is now accepted as the critical requirement for admission into the scientific marketplace.

Both proponents and critics have previously assumed, either tacitly or explicitly, that the outcome of a single investigation could be critical. Sidgwick believed that the results of the investigation of the Creery sisters were of this nature. The evidence was so strong, he argued, that the critics either had to now either accept the reality of telepathy or accuse the investigators of fraud [30]. Carpenter, rather than withhold judgment until independent investigators had either succeeded or failed in attempts to replicate Crookes' experiments with Home, acted as if he either had to agree to Crookes' claim or prove that Crookes had been duped. Both Price and Hansel insisted that it would be sufficient for Rhine and Soal to convince them of ESP if a parapsychology could perform successfully a single "fraud-proof" experiment.

The myth of the single, crucial experiment has resulted in needless controversy and has contributed to the False Dichotomy. Flew is just one who has argued convincingly that a single, unreplicated event which allegedly attests to a miracle, is simply a historical oddity which cannot be part of a scientific argument [3].

Apparently not all parapsychologists are convinced that the achievement of a repeatable psi experiment is either necessary or desirable for the advancement of parapsychology. The late J. G. Pratt argued that, "Psi is a spontaneous occurrence in nature, and we can no more predict precisely when it is going to occur in our carefully planned and rigorously controlled experiments than we can in everyday life psychic experiences.... Predictable repeatability is unattainable because of the nature of the phenomena" [69].

Pratt argued that parapsychology should give up the quest for the replicable experiment—an impossible goal in his opinion—and concentrate upon accumulating enough data on anomalous happenings to convince scientists and the public that psi is real. Other parapsychologists, however, realize that scientists are not going to be convinced until some semblance of replicability has been achieved. The late Gardner Murphy, while noting that replicability was not necessary for scientific acceptability in some areas of science, argued that for supporting claims for such irrational phenomena as psi, replicability was necessary. And, speaking as one of the dominant figures in parapsychology in 1971, he made it clear that he felt that parapsychology had a long way to go before it achieved replicable results [70].

Perhaps Honorton's position represents the contemporary position of the major parapsychologists [71]:

> Parapsychology will stand or fall on its ability to demonstrate replicable and conceptually meaningful findings. Future critics who are interested in the resolution rather than the perpetuation of the psi controversy are advised to focus their attention on systematic lines of research which are capable of producing such findings.

Psi and Repeatability

As the preceding quotation indicates, Honorton believes that critics should focus on "systematic lines of research" which apparently display replicable and/or "conceptually meaningful" findings. And, as we have seen, contemporary parapsychologists have offered us a number of such systematic lines to demonstrate that they have, in fact, already achieved the goals of repeatability and conceptual meaningfulness. The claims put forth in behalf of the altered state, random event generator, and remote viewing paradigms have already been cited. Similar claims have been made for work on correlates of psi such as attitudes and personality [72].

What can we expect if a critic, in an effort to be open-minded and responsible, accepts the challenge of Honorton and his fellow parapsychologists to examine the accumulated evidence from one or more of the "systematic lines" of inquiry? This challenge opens up a variety of possibilities. Which experiments should be included in the evaluation? It is impractical to consider all the experiments in parapsychology because even in this relatively sparsley populated area the number is by now enormous. In just considering a subset of experiments in the ESP area, Palmer, for example, covered approximately 700 experimental reports [72]. Including PK as well as ESP, I would estimate that, today, a determined critic, who wants to evaluate exhaustively all available experimental reports, might have to cope with upwards of 3000 experiments. Given my recent experience in trying to do justice to just 42 experiments on the Ganzfeld psi phenomenon [73], I would estimate that it could take a responsible critic over five years of almost full-time effort to properly evaluate this material.

Another problem facing both the proponent and critic is,

once a suitable sample of experiments has been selected, how to make an overall judgment about what patterns, trends, strengths, and weaknesses characterize the sample. Up until recently, such a review of a body of literature has been an unstructured and highly subjective affair. Understandably, two individuals surveying the same body of literature could, and did, often come up with diametrically opposed conclusions.

As cognitive psychologists have emphasized, the capacity of humans to handle mentally a number of items is severely limited. What constitutes an "item" varies greatly with the structure of the material and the individual's previous familiarity and expertise in a given field of knowledge. Even within his field of speciality, a scientist would have great difficulty in trying to comprehend patterns in over a dozen or so reports without external aids and a systematic procedure.

When the nonparapsychologist critic tries to make sense of a large body of parapsychological literature, he is at a great disadvantage. His critical capacities have not been trained to pick out relevant from irrelevant details in seeking interrelationships. Lacking concrete experience with many of the experimental designs, he is at a decided disadvantage in knowing what things could go wrong and which sorts of controls would be critical. And when the number of separate reports is more than a dozen or so, he cannot be expected to be able to grasp the total picture without help from systematic and quantitative summarization procedures.

Yet, so far as I can tell, only two critical evaluations of "systematic lines" of parapsychological research have ever been carried out with any procedure approximating systematic, explicit, and quantitative guidelines. Both of these were carried out fairly recently. One was by Charles Akers, a former parapsychologist with both experience and publications in the field [74]. The other was by myself, acting as an external critic who accepted the parapsychologists' challenge to fairly evaluate a systematic line of research which they feel represents their strongest case for the repeatable experiment [73], [75].

Akers' Methodological Criticisms of Parapsychology

Akers' methodological evaluation of contemporary parapsychological research represents a landmark in parapsychological criticism. Akers, who holds a Ph.D. degree in Social Psychology, has worked as a parapsychologist in Rhine's laboratory and knows the contemporary scene from the inside.

After a careful selection procedure, Akers arrived at a sample of 54 ESP experiments. These experiments had all been cited in the *Handbook of Parapsychology* or other parapsychological literature as exemplars of the evidential database. The selection was restricted to studies in which significant results had been claimed for a sample of relatively unselected percipients. He excluded unpublished reports, studies which were reported only as abstracts or convention reports, and studies which were exploratory or preliminary to a stronger replication. He also excluded experiments which produced scores in the wrong direction ("psi missing") [74].

> The final sample of 54 experiments is fairly complete. If it is not inclusive, it is at least representative of findings in altered state and personality research.

Akers then screened all his 54 studies sequentially through each of his several criteria to see how many could pass through all of them. He first looked at how many of the studies used inadequate randomization of the targets. Although he found almost half of the studies used inferior methods to randomize targets he considered this to be a "minor contaminant." In his opinion, such randomization failures as he observed would not be sufficient to account for the above chance results which each of these studies obtained.

Next he looked at the possibility of sensory leakage. For example, in several of the Ganzfeld experiments the agent handled the slide or picture which served as the target. Later the percipient was given that very same target along with some foils and asked to select which item had been the target. In such a situation either inadvertent or deliberate cueing is clearly a possibility. A parapsychologist should not be entitled to claim ESP as the explanation for a successful selection by the percipient under such circumstances. Akers assigned a flaw to any experiment which had this or one of his other categories of possibilities for sensory leakage. As many as 22 of the 54 experiments were cited for having at least one flaw of the sensory leakage kind (some had more than one kind).

In a similar fashion, Akers checked for security problems, recording errors, optional stopping, data selection, inadequate documentation, multiple testing, and some additional flaws of a technical nature. On each criterion, Akers assigned a flaw only if, in his opinion, the defect was sufficient to account for the above chance hitting actually reported [74].

> Results from the 54-experiment survey have demonstrated that there are many alternative explanations for ESP phenomena; the choice is not simply between psi and experimenter fraud.... The numbers of experiments flawed on various grounds were as follows: randomization failures (13), sensory leakage (22), subject cheating (12), recording errors (10), classification or scoring errors (9), statistical errors (12), reporting failures (10).... All told, 85 percent of the experiments were considered flawed (46/54).

In other words, only 8 of the 54 experiments—all of which were selected to be best cases—were free of at least one serious flaw on Akers' criteria. But Akers points out a number of reasons to be concerned about the adequacy of even these "flawless" studies [74].

> In conclusion, there were eight experiments conducted with reasonable care, but none of these could be considered as methodologically strong. When all 54 experiments are considered, it can be stated that the research methods are too weak to establish the existence of a paranormal phenomenon.

Akers' conclusion is especially damaging to the case for psi because he leaned over backwards to give the benefit of doubt to the experimenters. In some cases where the documentation was incomplete, Akers assumed that the investigator had taken the proper precautions against sensory leakage. And Akers did not assign flaws to experiments if their randomization procedures were less than optimal (he considered this to be only a "minor contaminant"). Experiments that were deficient on his other criteria such as optional stopping and others were not assigned flaws if, on Akers' judgment, the deficiency on that criterion was insufficient to have caused the total number of hits. In other words, Akers was not judging whether the experiment had

met standards of scientific acceptability, but rather, he was assigning flaws if a given deficiency *by itself* was sufficient to have accounted for the results. And, finally, Akers did not consider the possibility that *combinations* of deficiencies, each in themselves being insufficient, might have been more than enough to account for the reported findings.

Hyman's Critique of the Ganzfeld Experiments

Although Akers' and my critiques were conducted independently, and although our samples and procedures differed in many important ways, we came to essentially the same conclusion. In spite of claims for both scientific confirmation of psi and repeatibility within certain systematic lines of research, both Akers and I concluded that the best contemporary research in parapsychology does not survive serious and careful scientific scrutiny. Parapsychology is not yet ready to bring its case before the general scientific public.

My approach was to look for a research program in parapsychology that consisted of a series of experiments by a variety of investigators and that was considered by parapsychologists as especially promising. I quickly discovered a systematic body of research which many of the leading parapsychologists considered to be the most promising one on the contemporary scene. This research program was based on the Ganzfeld/psi paradigm.

The word "Ganzfeld" is German for total field. It is used to describe a technique in the study of perception which creates a visual field with no inhomogeneities. The motivation for creating such a visual field stems from certain theoretical predictions of Gestalt psychology. A recently developed and simple procedure for creating such a Ganzfeld is to tape halves of ping pong balls over the eyes of subjects. A bright light is then directed to the covered eyes. The percipient experiences a visual field with no discontinuities and describes the perceptual effect as like being in a fog.

The parapsychologists became interested in the Ganzfeld when it was reported that subjects who experience the Ganzfeld quickly enter into a pleasant, altered state. They adopted it as a quick and easy way to place percipients into a state that they felt would be conducive to the reception of psi signals. In a typical Ganzfeld/psi experiment, the percipient has the pin pong balls taped over his eyes and then is placed in a comfortable chair or reclines on a bed. In addition to a bright light shining on the halved ping pong balls, white noise or the sound of ocean surf is fed into the percipient's ears through earphones.

After 15 min or so in this situation, the percipient is presumed ready to receive the psi signal. An agent, in another room or building, is given a target which is randomly selected from a small pool, say, of four pictures (the pool of pictures has been selected, in turn, by random means from a large collection of such pools). The agent concentrates or studies the target during a predetermined time interval. At the same time the percipient, isolated in a relatively sound-proofed chamber, freely describes all the associations and impressions that occur to him during the sending interval.

At the end of the session the halved ping pong balls are removed. The pool of pictures for that trial, including the target, are brought to the percipient. The percipient then indicates, by ranking or rating, how close each of the items in the pool are to the impressions that occurred to him or her during the Ganzfeld session. The most typical scoring procedure classifies the outcome as a "hit" if the percipient correctly judges the actual target as closest to the Ganzfeld impressions.

In the typical experiment a pool of four target candidates is used on each trial. Over a number of trials, the percipients would be expected to achieve hits on 25 percent of the trials just by chance. If the actual rate of hitting is significantly above this chance level, then it is assumed, given that proper experimental controls have been employed, that ESP has probably operated.

Charles Honorton, the parapsychologist who first published a Ganzfeld/psi experiment [76] and who also has strongly defended the paradigm as "psi conducive," responded to my request for cooperation by undertaking to supply me with copies of every relevant report between 1974—the date of the first published Ganzfeld/psi experiment—and the end of 1981—the year I made the request. In January 1982 I received a package containing 600 pages of reports on the Ganzfeld/psi experiment.

The experiments in the database given to me for examination were extracted from 34 separate reports written or published from 1974 through 1981. By Honorton's count, these 34 reports described 42 separate experiments. Of these, he classified 23 as having achieved overall significance on the primary measure of psi at the 0.05 level. This successful replication rate of 55 percent is consistent with earlier estimates of success for this paradigm which ranged from 50 to 58 percent [73]. Approximately half of these experiments had been published in refereed journals or monographs. The remainder had appeared only as abstracts or papers delivered at meetings of the Parapsychological Association. The studies had been authored by 47 different investigators, many of them prominent members of the Parpsychological Association.

The details of my analysis and my conclusions have been published in the *Journal of Parapsychology* [73]. The same issue of that journal contains Honorton's detailed rebuttal to my critique [77]. Here I will merely supply the barebones of my critique.

1) I first examined the claim that the proportion of successful replications of the Ganzfeld/psi experiment was 55 percent. This estimate, it turned out, was based upon a number of questionable assumptions. Much ambiguity exists as to what the unit of analysis should be. In some cases, the individual experimental conditions within a single complicated experiment were each counted as separate "experiments." In other cases, the pooled data over a number of separate experimental conditions were counted as a single unit. That this can make a difference is shown by the fact that when I tried to apply a consistent criterion to the database for determining individual units, I came up with a success rate closer to 30 than to 50 percent. Other considerations such as unknown experiments lead me to conclude that the actual success rate, defining "success" according to Honorton's criterion, was probably around 30 percent.

2) But even a success rate of 30 percent is impressive if the actual rate of success to be expected by chance was the assumed 5 percent. I pointed to a variety of examples in which multiple tests were applied to the same data in such

a way as to inflate the actual probability for success just by chance over the assumed rate. Taking into consideration a number of factors, I estimated that the actual chance level could easily be 25 percent or higher.

3) In addition to analyses that inadvertently inflated the significance levels, I noted a number of other departures from optimal experimental procedure that could have artificially contributed to the outcomes. These flaws could be clustered into three categories: Security, Statistical, and Procedural. Security flaws included failure to preclude sensory cues as well as loose monitoring of critical aspects in the experiment. Statistical flaws consisted of wrong use of statistical procedures. Procedural flaws consisted of inadequate randomization of targets, incomplete documentation, and possible problems at feedback. What was both surprising and dismaying to me was that not a single experiment in the database was free from at least one of these defects. These defects were chosen to be those that I assume most parapsychologists would agree should not be part of a well-conducted experiment.

4) I tried to make it clear that I was not assuming that these flaws were the cause of the observed results. Rather, I assumed that the presence of such defects could be taken as a symptom that the experiment had not been conducted with adequate care. Indeed, it was clear that at least some of the experiments in the database had been intended to serve only as pilot or preliminary experiments. Nevertheless, I did look at the correlation between the three clusters and success of the experiment. Although the Security and the Statistical clusters did not correlate with outcome, the Procedural cluster did correlate with the probability of obtaining a significant outcome. Honorton strongly disagrees with this conclusion [77].

As a result of my detailed examination of the claims for the Ganzfeld/psi findings, I concluded my long report as follows [73]:

> In conclusion, the current data base has too many problems to be seriously put before outsiders as evidence for psi. The types of problems exhibited by this data base, however, suggest interesting challenges for the parapsychological community. I would hope that both parapsychologists and critics would wish to have parapsychological experiments conducted according to highest standards possible. If one goal is to convince the rest of the scientific community that the parapsychologists can produce data of the highest quality, then it would be a terrible mistake to employ the current Ganzfeld/psi data base for this purpose. Perhaps the Parapsychological Association can lead the way by setting down guidelines as to what should constitute an adequate confirmatory experiment. And, then, when a sufficient number of studies have accumulated which meet these guidelines, they can be presented to the rest of the scientific community as an example of what parapsychology, at its best, can achieve. If studies carried out according to these guidelines also continue to yield results suggestive of psi, then the outside scientific community should be obliged to take notice.

Honorton, not surprisingly, disagrees with my conclusions [77]. After my critique was completed, Honorton carried out a revised and different analysis of the database. He claims his new analysis eliminates my criticisms about inflated significance levels. Honorton also developed his own [illegible] for evaluating the methodological quality of each [illegible]. According to his ratings, there is no correlation [illegible] the experiment and its outcome.

The problem that both of us face when judging the quality of the individual experiments is that we are doing this after the fact. Although we agree on several of our ratings, we tend to disagree in ways which suggest our presumed biases. Honorton tends to find more defects in the unsuccessful experiments than I do. On the other hand, I tend to find more defects in the successful experiments than Honorton does. In the absence of double-blind ratings, this aspect of our disagreement represents a stalemate.

However, whether one uses Honorton's or my ratings, the number of departures from accepted methodological procedure is unacceptably high for this database. Although Honorton and I disagree on whether the observed flaws weaken the case for psi, we do not disagree that they exist. So far as I can tell, no parapsychologist has provided an explanation of why almost all of the experiments in this database have at least one of these flaws.

Conclusions

With the exception of the contemporary parapsychological literature, the evidence for psi reviewed in this paper comes from investigations which today's parapsychologists would not put before us as part of their strongest case for psi. Many of these parapsychologists might believe I was being unfair in dwelling upon these castoffs from the past. But it is just this fact that the cases I have examined are now castoffs which brings up important questions about how to approach the contemporary evidence.

Each of the cases from the past which I have discussed were, in their own time, considered to be by the parapsychologists of that day examples of scientifically sound evidence for psi. It is only subsequent generations, for the most part, who have set the preceding exemplars aside. In some cases the reasons for the abandonment of what was once a foundation stone in the case for psi are clear. Subsequent investigators or critics found previously unrecognized defects in the studies or strong suspicions of fraud had been generated. Other experimental paradigms have disappeared from the database for less obvious reasons.

Some previously successful paradigms have disappeared because they no longer seem to yield significant results. Others such as the sheep–goats design seem to have simply gone out of fashion. One major parapsychologist once told me that it seems to be the ultimate fate of every successful paradigm to eventually lose its ability to yield significant results. He believed this was related to the fact that psi depends both upon the novelty of the design and the motivations of the experimenter. At first a new paradigm generates excitement and optimism. But after it has been around for a while, the initial excitement and enthusiasm abates and the experimenter no longer communicates the original emotions that accompanied the paradigm when it was still relatively new.

But, whatever the reason, each generation's best case for psi is cast aside by subsequent generations of parapsychologists and are replaced with newer, more up-to-date best cases. Not only does the evidence for psi lack replicability, but, unlike the evidence from other sciences, it is noncumulative. It is as if each new generation wipes the slate clean and begins all over again. Consequently, the eviden-

tial database for psi is always shifting. Earlier cases are dropped and replaced with newer and seemingly more promising lines of research. [One of the readers of this paper argues that it is only partially true that parapsychological research is noncumulative. Although his argument might have some validity, I do not think it changes the point I am making here.]

The late J. G. Pratt, in challenging his parapsychological colleagues' hopes for a repeatable experiment, wrote [69]:

> One could almost pick a date at random since 1882 and find in the literature that someone somewhere had recently obtained results described in terms implying that others should be able to confirm the findings. Among those persons or groups reflecting such enthusiasm are the S.P.R. Committee on Thought-Transference; Richard Hogson (in his investigation of Mrs. Piper); Feilding, Baggally, and Carrington (in their Palladino investigations); J. B. Rhine (work reported in *Extra-Sensory Perception*); Whately Carington (in his work on paranormal cognition of drawings); Gertrude Schmeidler (in her sheep-goat work); Van Bussbach, and Anderson and White (in their research on teacher-pupil attitudes); the Maimonedes dream studies; the Stepanek investigators; the investigators of Kulagina's directly-observable PK effects; research using the ganzfeld technique; and the SRI investigators ("remote viewing"). One after another, however, the specific ways of working used in these initially successful psi projects have fallen out of favor and faded from the research scene—except for the latest investigations which, one may reasonably suppose, have not yet had enough time to falter and fade away as others before them have done.

When Pratt wrote those words in 1978, the "latest investigations" included the Ganzfeld/psi experiments, the Remote Viewing investigations, and the PK research using Random Event Generators. These would have been among the contemporary investigations which, given Pratt's pessimistic extrapolations, "one may reasonably suppose, have not yet had enough time to falter and fade away as others before them have done." Today, signs do seem to indicate that these seemingly "successful" lines of research may be much weaker than had been previously advertised [24], [74], [75].

However, as always, new and more promising lines of work seem to be ready to take their place. Honorton and his colleagues at the Psychophysical Research Laboratories in Princeton, NJ, seem to be developing a number of very promising lines of research [78]. They have been developing a completely automated version of the Ganzfeld experiment which eliminates many of the problems raised by my critique. They have also been perfecting a "transportable" experiment—one that can be carried out by any investigator who has access to an Apple personal computer. The experiment, also completely automated, is a variation of the Random Event Generator paradigm but with a variety of built-in safeguards which apparently eliminate almost all the options for multiple testing.

Nearby, but completely independent of the work going on at the Psychophysical Laboratories, is the research on anomalous phenomena being carried out by Robert Jahn and his associates in the School of Engineering and Applied Science at Princeton University [1], [79], [80]. For more than five years Jahn and his associates have been perfecting the instrumentation and experimental designs for conducting sophisticated variations of both the remote viewing paradigm and the PK work with random event generators. Although they have collected large databases for each of these paradigms, most of the work has been reported only in technical reports. The reported findings do seem impressive, but they have yet to be described in sufficient detail for a full-scale evaluation. And, given both the scale of the effort and the sophistication of the methodology and instrumentation, it will be many years before adequate replications in independent laboratories will be possible.

As promising as this most recent work by Honorton and Jahn might seem to be, none of it has reached a stage where it is ready for a full-scale critical evaluation. Already, the sharp-eyed critic can detect both inconsistencies with previous findings in the same lines of research and departures from ideal practice. As the history of parapsychology teaches us, we will have to wait for several more years before we can adequately judge if somehow these latest efforts can avoid the fate that all their promising predecessors have suffered.

Perhaps, however, history does not have to repeat itself in all its depressing aspects. And I can see some encouraging signs of breaks with previous patterns in the way proponents carry out and defend their findings and the way critics respond.

Since its inception as an institutionalized undertaking, psychical research has suffered from the lack of relevant, informed, and constructive criticism. This particular deficiency seems to be changing. For one thing, the younger generation of parapsychologists have produced some internal critics who are both knowledgeable and effective. In addition to Akers, there are others such as Susan Blackmore, Adrian Parker, Gerd Hövelmann, and J. E. Kennedy who have recognized the current deficiencies of parapsychological research and have a strong committment to raising the standards. Although it is still difficult to find external critics who are both informed and constructive, one can see some indications that this situation may also improve.

Another positive sign is the attempt to replace subjective, impressionistic evaluations of the parapsychological literature with more systematic, explicit assessments. Both Honorton [77] and I [73] have used "meta-analysis" in our dispute over the adequacy of the Ganzfeld/psi database. "Meta-analysis" is a term coined to describe the approach to reviewing a body of research which makes the various phases as explicit and quantitative as feasible [81], [82].

> The approach to research integration referred to as "meta-analysis" is nothing more than the attitude of data analysis applied to quantitative summaries of individual experiments. By recording the properties of studies and their findings in quantitative terms, the meta-analysis of research invites one who would integrate numerous and diverse findings to apply the full power of statistical methods to the task. Thus it is not a technique; rather it is a perspective that uses many techniques of measurement and statistical analysis.

(From [81].)

Meta-analysis is by no means a panacea. Much subjectivity remains on such matters as which studies to include and exclude from the sample, how to score the "effect size" or degree of success of a study, what variables to include, how to assign studies values on the variables, and what should be the sampling unit. In addition, many serious problems have to be resolved about how to cope with the fact that individual studies are not independent and the analyses are

conducted "*post hoc*." Yet, it has many advantages over the previously unstructured and subjective assessments. The reviewer is forced to make many more of his or her standards and procedures explicit. The resulting debate can be more focussed and the specific areas of disagreement can be pinpointed more accurately. In addition, the use of quantitative summaries often brings out patterns and relationships that would ordinarily escape the unaided reviewer's cognitive limits.

Along with an increase in more informed and constructive criticism there are signs that the parapsychological community is responsive and willing to change both its procedures and claims in line with some of the criticisms. Although we still disagree strongly on many of the issues, Honorton has made many changes in his claims and procedures in a sincere effort to take some of my criticisms into account [73], [77]. At its 1984 annual meetings in Dallas, TX, the Parapsychological Association established a committee which will attempt to establish guidelines for the performance of acceptable experiments in various lines of parapsychological research. Along with some major parapsychologists such as Honorton, the committee includes both internal critics such as Akers and external ones such as myself.

My survey of psychical research from the time of Hare and Crookes to the present has suggested that, although the specific evidence put forth to support the existence of psi changes over time, many of the key issues and controversies have remained unchanged. The parapsychologists still employ similar strategems to seemingly enable them to stick to their claims in the face of various incosistencies. And the critics, sharing many assumptions with the proponents, still behave in rather emotional and irrational ways. Indeed, the level of the debate during the preceding 130 years has been an embarrassment for anyone who would like to believe that scholars and scientists adhere to standards of rationality and fair play.

I suspect it is because the quality of the criticism has been so poor and its content so obviously irrelevant that parapsychologists have managed to live so long with the illusion that the quality of their evidence was so much better than it really was. Both Akers and I were surprised to find how defective, in terms of the most elementary standards, the best of the contemporary parapsychological research really was. I know that some parapsychologists have been surprised to realize how far the current status of psi research departs from the professed standards of their field. And I would not be surprised that most of the rest of the parapsychological community, in the absence of systematic and critical surveys, had assumed that their database was of a much higher quality than it, in fact, is.

All this suggests, as I have already indicated, that the parapsychological evidence, despite a history of more than 130 years of inquiry, is not ready to be placed before the scientific community for judgment. The parapsychologists' first order of business should be to get their own house in order. They no longer can safely assume that the typical parapsychologist has the competence to correctly use statistical tools, design appropriate investigations, carry out these investigations correctly, or to write them up properly. Indeed, the evidence suggests the opposite. Both the Parapsychological Association and the parapsychological journals have to establish explicit guidelines and minimal standards. Then they have to make sure that members of their profession become fully aware of these standards and recognize the necessity for living up to them.

References

[1] R. G. Jahn, "The persistent paradox of psychic phenomena: an engineering perspective," *Proc. IEEE*, vol. 70, pp. 136–170, Feb. 1982.

[2] G. Murphy and R. O. Ballou, *William James on Psychical Research*. New York: Viking Press, 1969.

[3] A. Flew, "Parapsychology revisited: laws, miracles and repeatability," in *Philosophy and Parapsychology*, J. Ludwig, Ed. Buffalo, NY: Prometheus Books, pp. 263–269.

[4] R. Hyman, "Pathological science: towards a proper diagnosis and remedy," *Zetetic Scholar*, no. 6, pp. 31–39, July 1980.

[5] R. L. Moore, *In Search of White Crows: Spiritualism, Parapsychology and American Culture*. New York: Oxford Univ. Press, 1977.

[6] F. Podmore, *Mediums of the 19th Century*. New Hyde Park NY: University Books, 1963 (two volumes).

[7] M. Faraday, "Experimental investigation of table-moving," *The Athenaeum*, pp. 801–803, July 2, 1853.

[8] R. Hare, *Experimental Investigation of the Spirit Manifestations, Demonstrating and Existence of Spirits and their Communion with Mortals: Doctrine of the Spirit World Respecting Heaven, Hell, Morality, and God. Also, the Influence of Scripture on the Morals of Christians*. New York: Partridge & Brittan, 1855.

[9] L. Shepard, Ed., *Encyclopedia of Occultism and Parapsychology*. Detroit, MI: Gale Research Co., 1978.

[10] I. Asimov, *Asimov's Biographical Encyclopedia of Science and Technology*. New York: Equinox, 1976.

[11] W. George, *Biologist Philosopher: A Study of the Life and Writings of Alfred Russel Wallace*. New York: Abelard-Schuman, 1964.

[12] M. J. Kottler, "Alfred Russel Wallace, the origin of man, and spiritualism," *Isis*, vol. 65, pp. 145–192, 1974.

[13] F. M. Turner, *Between Science and Religion: The Reaction to Scientific Naturalism in Late Victorian England*. New Haven, CT: Yale Univ. Press, 1974.

[14] A. R. Wallace, *My Life: A Record of Events and Opinions*. New York: Dodd, Mead, & Co., 1906.

[15] R. G. Medhurst, Ed., *Crookes and the Spirit World: A Collection of Writings by or Concerning the Work of Sir William Crookes, O.M., F.R.S., in the Field of Psychical Research*. New York: Taplinger, 1972.

[16] E. E. F. D'Albe, *The Life of Sir William Crookes*. New York: D. Appleton and Co., 1924.

[17] J. Palfreman, "William Crookes: spiritualism and science," *Ethics Sci. Med.*, vol. 3, pp. 211–227, 1976.

[18] S. J. Davey, "The possibilities of mal-observation and lapse of memory from a practical point of view: Experimental investigation," *Proc. Soc. Psychical Res.*, vol. 4, pp. 405–495, 1887.

[19] E. J. Dingwall, *Very Peculiar People: Portrait Studies in the Queer, the Abnormal and the Uncanny*. New Hyde Park, NY: University Books, 1962.

[20] J. Hasted, *The Metal-Benders*. London, England: Routledge & Kegan Paul Ltd., 1981.

[21] R. Hyman, [Review of *The Geller Papers*], *The Zetetic* [*The Skeptical Inquirer*], vol. 1, pp. 73–80, Fall/Winter 1976.

[22] ———, "Psychics and scientists: 'Mind-Reach' and remote viewing," *The Humanist*, vol. 37, pp. 16–20, May/June 1977.

[23] ———, "Scientists and psychics," in *Science and the Paranormal*, G. O. Abell and B. Singer, Eds. New York: Scribner's, 1981, pp. 119–141.

[24] ———, "Outracing the evidence: the muddled 'Mind Race'," *The Skeptical Inquirer*, vol. 9, pp. 125–145, 1984–1985.

[25] D. Marks and R. Kammann, *The Psychology of the Psychic*. Buffalo, NY: Prometheous Books, 1980.

[26] J. Palfreman, "Between scepticism and credulity: A study of Victorian scientific attitudes to modern spiritualism," in *On*

the Margins of Science: The Social Construction of Rejected Knowledge, R. Wallis, Ed. Staffordshire, England: Univ. of Keele, 1979, pp. 201–246.

[27] C. Panati, Ed. *The Geller Papers: Scientific Observations on the Paranormal Powers of Uri Geller*. Boston, MA: Houghton Mifflin, 1976.

[28] R. Targ and H. E. Puthoff, *Mind Reach: Scientists Look at Psychic Abilities*. New York: Delacorte, 1977.

[29] J. C. F. Zöllner, *Transcendental Physics*. New York: Arno Press, 1976 (Reprint of 1888 edition published by Colby & Rich, Boston).

[30] H. Sidgwick, [Presidential Address], *Proc. Soc. Psychical Res.*, vol. 1, pp. 7–12, 1882–1883.

[31] J. B. Hasted, D. Bohm, E. W. Bastin, and B. O'Regan, "Experiments on psychokinetic phenomena," in *The Geller Papers*, C. Panati, Ed. Boston, MA: Houghton Mifflin, 1976, pp. 183–196.

[32] A. Gauld, *The Founders of Psychical Research*. New York: Schocken, 1968.

[33] W. F. Barrett, E. Gurney, and F. W. H. Myers, "First report on thought-reading," *Proc. Soc. Psychical Res.*, vol. 1, pp. 13–34, 1882–1883.

[34] E. Gurney, F. W. H. Myers, and W. F. Barrett, "Second report on thought-transference," *Proc. Soc. Psychical Res.*, vol. 1, pp. 70–89, 1882–1883.

[35] E. Gurney, F. W. H. Myers, F. Podmore, and W. F. Barrett, "Third report on thought-transference," *Proc. Soc. Psychical Res.*, vol. 1, pp. 161–181, 1882–1883.

[36] F. Nicol, "The founders of the S. P. R.," *Proc. Soc. Psychical Res.*, vol. 55, pp. 341–367, 1972.

[37] W. F. Barrett, E. Gurney, and F. W. H. Myers, "Thought-reading," *The Nineteenth Century*, vol. 11, pp. 890–900, June 1882.

[38] J. E. Coover, *Experiments in Psychical Research*. Stanford, CA: Stanford Univ. Press, 1917.

[39] E. Gurney, F. W. H. Myers, and F. Podmore, *Phantasms of the Living*. London, England: Trubner's, 1886.

[40] G. S. Hall, [Reviews of the *Proceedings of the English Society for Psychical Research*, July, 1882 to May 1887, and of *Phantasms of the Living*], *Amer. J. Psychol.*, vol. 1, pp. 128–146, 1887–1888.

[41] T. H. Hall, *The Strange Case of Edmund Gurney*. London, England: Duckworth, 1964.

[42] O. Pfungst, *Clever Hans*. New York: Holt, 1911.

[43] R. N. Giere, *Understanding Scientific Reasoning*. New York: Holt, Rinehart & Winston, 1979.

[44] E. Gurney, "Note relating to some published experiments in thought-transference," *Proc. Soc. Psychical Res.*, vol. 5, pp. 269–270, 1898.

[45] C. Tart, [Review of *The Magic of Uri Geller* by the Amazing Randi; *The Geller Papers* edited by Charles Panati; and *My Story* by Uri Geller], *Psychol. Today*, pp. 93–94, July 1976.

[46] J. B. Rhine, *Extra-Sensory Perception*. Boston, MA: Bruce Humphries, 1935 [Originally published in 1934 by the Boston Society for Psychic Research].

[47] S. H. Mauskopf and M. R. McVaugh, *The Elusive Science: Origins of Experimental Psychical Research*. Baltimore, MD: The Johns Hopkins Press, 1980.

[48] B. Wolman, Ed., *Handbook of Parapsychology*. New York: Van Nostrand-Reinhold, 1977.

[49] K. M. Goldney, "The Soal-Goldney experiments with Basil Shackleton (BS): A personal account," *Proc. Soc. Psychical Res.*, vol. 56, pp. 73–84, 1974.

[50] S. G. Soal and F. Bateman, *Modern Experiments in Telepathy*, 2nd. ed. New Haven, CT: Yale Univ. Press, 1954.

[51] K. M. Goldney, "Obituary: Dr. S. G. Soal, M. A., D. Sc.," *J. Soc. Psychical Res.*, vol. 48, pp. 95–98, 1975.

[52] R. A. McConnell, [Review of *Modern Experiments in Telepathy*], *J. Parapsychol.*, vol. 18, pp. 245–258, 1954.

[53] G. R. Price, "Science and the supernatural," *Science*, vol. 122, pp. 359–367, 1955.

[54] S. G. Soal, "On 'Science and the supernatural,'" *Science*, vol. 123, pp. 9–11, 1956.

[55] J. B. Rhine, "Comments on 'Science and the supernatural,'" *Science*, vol. 123, pp. 11–14, 1956.

[56] C. Scott and P. Haskell, "Fresh light on the Shackleton experiments?" *Proc. Soc. Psychical Res.*, vol. 56, pp. 43–72, 1974.

[57] R. C. Medhurst, "The fraudulent experimenter: Professor Hansel's case against psychical research," *J. Soc. Psychical Res.*, vol. 44, pp. 217–243, 1968.

[58] C. E. M. Hansel, *ESP: A Scientific Evaluation*. New York: Scribner's, 1966.

[59] ———, *ESP and Parapsychology: A Critical Re-Evaluation*. Buffalo, NY: Prometheus Books, 1980.

[60] B. Markwick, "The Soal-Goldney experiments with Basil Shackleton: New evidence of data manipulation," *Proc. Soc. Psychical Res.*, vol. 56, pp. 250–277, 1978.

[61] W. Broad and N. Wade, *Betrayers of the Truth*. New York: Simon & Schuster, 1982.

[62] J. Hixson, *The Patchwork Mouse*. New York: Doubleday, 1976.

[63] T. Hall, "Florence Cook and William Crookes: A footnote to an enquiry," *Tomorrow*, vol. 11, pp. 341–359, Autumn 1963.

[64] C. Honorton, "Psi and internal attention states: Information retrieval in the ganzfeld," in *Psi and States of Awareness*, B. Shapin and L. Coly, Eds. New York: Parapsychology Foundation, 1978, pp. 79–90.

[65] ———, "Psi and internal attention states," in *Handbook of Parapsychology*, B. B. Wolman, Ed. New York: Van Nostrand-Reinhold, 1977, pp. 435–472.

[66] E. C. May, B. S. Humphrey, and G. S. Humbbard, "Electronic system perturbation techniques," SRI Int., Final Rep., Sept. 30, 1980.

[67] H. E. Puthoff and R. Targ, "A perceptual channel for information transfer over kilometer distances: Historical perspective and recent research," *Proc. IEEE*, vol. 64, pp. 329–354, 1976.

[68] R. Targ and K. Harary, *The Mind Race: Understanding and Using Psychic Abilities*. New York: Villard Books, 1984.

[69] J. G. Pratt, "Prologue to a debate: Some assumptions relevant to research in parapsychology," *J. Amer. Soc. Psychical Res.*, vol. 72, pp. 127–139, 1978.

[70] G. Murphy, "The problem of repeatability in psychical research," *J. Amer. Soc. Psychical Res.*, vol. 65, pp. 3–16, 1971.

[71] C. Honorton, "Beyond the reach of sense: Some comments on C. E. M. Hansel's *ESP and Parapsychology: A Critical Re-Evaluation*," *J. Amer. Soc. Psychical Res.*, vol. 75, pp. 155–166, 1981.

[72] J. Palmer, "Extrasensory perception: Research findings," in *Advances in Parapsychological Research, vol. 2: Extrasensory Perception*, S. Krippner, Ed. New York: Plenum, 1978, pp. 59–243.

[73] R. Hyman, "The Ganzfeld Psi experiment: A critical appraisal," *J. Parapsychol.*, vol. 30, pp. 76–82, 1985.

[74] C. Akers, "Methodological criticisms of parapsychology," in *Advances in Parapsychological Research*, vol. 4, S. Krippner, Ed. Jefferson, NC: McFarland, 1984.

[75] R. Hyman, "Does the Ganzfeld experiment answer the critics' objections?" in *Research in Parapsychology*, W. G. Roll, J. Beloff, and R. A. White, Eds. Metuchen, NJ: Scarecrow Press, 1983, pp. 21–23.

[76] C. Honorton and S. Harper, "Psi-mediated imagery and ideation in an experimental procedure for regulating perceptual input," *J. Amer. Soc. Psychical Res.*, vol. 68, pp. 156–168, 1974.

[77] C. Honorton, "Meta-analysis of psi ganzfeld research: A response to Hyman," *J. Parapsychol.*, vol. 49, pp. 51–91, 1985.

[78] Psychophysical Research Laboratories, *1983 Annual Report*, Princeton, NJ.

[79] R. D. Nelson, B. J. Dunne, and R. G. Jahn, "An REG experiment with large data base capability, III: Operator related anomalies," Tech. Note, School of Eng./Appl. Sci. Princeton Univ., Princeton, NJ, Sept. 1984.

[80] B. J. Dunne, R. G. Jahn, and R. D. Nelson, "Precognitive remote perception," Tech. Note, School of Eng./Appl. Sci., Princeton Univ., Princeton, NJ, Aug. 1983.

[81] G. V. Glass, B. McGaw, and M. L. Smith, *Meta-Analysis in Social Research*. Beverly Hills, CA: Sage Publ., 1981.

[82] L. V. Hedges and I. Olkin, "Analyses, reanalyses, and meta-analysis [Review of *Meta-Analysis in Social Research*]," *Contemp. Educ. Rev.*, vol. 1, pp. 157–165, 1982.