

Does ESP Exist?

Daryl J. Bem

Daryl J. Bem is professor of psychology at Cornell University. He became involved in the world of psi research in 1983, when he was asked to critically examine the experimental procedures for the autoganzfeld studies from the perspective of an experimental psychologist and mentalist (a magician who specializes in the simulation of psi).

Reports of psychic phenomena are as old as human history. Experimental tests of psychic phenomena are almost as old. According to Herodotus, the ancient Greek historian, King Croesus of Lydia dispatched several of his men to test seven oracles to see if any of them could divine what he, the king, was doing on the day of the test. Only Pythia, priestess of Apollo at Delphi, was able to divine correctly that the king was making a lamb and tortoise stew in a bronze kettle.

Convinced of her powers, Croesus then posed the question that really interested him: If he attacked the rival kingdom of Persia, would he be able to defeat its army? Pythia replied, "When Croesus has the Halys crossed, a mighty empire will be lost." Insufficiently alert to the ambiguity of this prediction, Croesus crossed the river, attacked, and lost his mighty empire. Evaluating "psychic" data is a risky business.

The contemporary technical term for psychic phenomena is psi. More precisely, psi denotes anomalous processes of information or energy transfer, processes that are currently unexplained in terms of known physical or biological mechanisms. These processes

Recent laboratory research suggests that parapsychologists might finally have cornered their elusive quarry: reproducible evidence for psychic functioning.

include *extrasensory perception* (ESP), the acquisition of information without using the known senses, and *psychokinesis*, the ability to affect physical objects or events without the intervention of any known physical force.

In turn, ESP comprises the following:

- *Telepathy*. The transfer of information from one person to another without the mediation of any known channel of sensory communication.

- *Clairvoyance*. The acquisition of information about places, objects, or events without the mediation of any of the known senses (for example, Pythia's knowledge that the king was making stew).

- *Precognition*. The acquisition of information about a future event that could not be anticipated through any known inferential process. (Pythia's prediction about the loss of an empire, although dubious, is an example.)

Serious scholarly investigation of psi began in 1892, when a group of scholars in London founded the Society for Psychical Research (SPR) to

investigate that large body of debatable phenomena designated by such terms as mesmeric, psychical and spiritualistic . . . without prejudice or prepossession of any kind, and in the same spirit of exact and unimpassioned inquiry which has enabled Science to solve so many problems, once not less obscure nor less hotly debated.

The SPR was active until the early years of the twentieth century, when many of the original founders had died and enthusiasm declined.

Contemporary psi research is usually considered to have begun in 1927, when Joseph Banks Rhine and his wife/collaborator, Louisa, arrived in the psychology department at Duke University in Durham, North Carolina. Rhine's experiments, which tested for ESP with decks of cards containing geometric symbols, became well known to the general public in 1937 when he published *New Frontiers of the Mind*. The book received widespread press coverage and became a Book of the Month Club selection. Even today, many Americans know of Rhine's work.

Since Rhine, many parapsychologists have reported positive psi results using a wide variety of experimental procedures. Yet most academic psychologists are not yet persuaded that the existence of psi has been established.

Searching for a repeatable experiment

In science generally, a phenomenon is not considered established until it has been observed repeatedly by several researchers. This criterion has been the source of the most serious criticism of parapsychology: that it has failed to

word meaning "total field." To quiet "noise" produced by internal bodily tension, the receiver is also led through a set of relaxation exercises at the beginning of the ganzfeld period.

While the receiver is in the ganzfeld, a second subject (the sender) sits in a separate soundproof room and concentrates on the "target," a ran-

ganzfeld researcher. At that time, there had been 42 reported ganzfeld studies conducted by investigators in 10 laboratories.

Across these studies, receivers achieved an average hit rate of about 35 percent. (This might seem like a small margin of success over the 25 percent hit rate expected by chance, but a person with this margin of advantage in a gambling casino would get rich very quickly.) Statistically this result is highly significant: The odds against getting a 35 percent hit rate across that many studies by chance are greater than a billion to one.

By the late 1960s, several parapsychologists had become dissatisfied with the repetitive forced-choice procedures pioneered by Rhine.

yield a single reliable demonstration of psi that can be replicated by other investigators. In 1974, an experimental procedure was introduced that holds out the promise of supplying that repeatable demonstration: the ganzfeld procedure.

By the late 1960s, several parapsychologists had become dissatisfied with the repetitive forced-choice procedures pioneered by Rhine, believing that they failed to capture the kinds of psi experiences that people report in everyday life. Both historically and cross-culturally, psi has usually been associated with dreaming, meditation, trances of various kinds, and other altered states of consciousness. This suggested that psi information may function like a weak signal normally masked by the sensory "noise" of everyday life. Thus, altered states of consciousness may enhance a person's ability to detect psi information simply because they reduce interfering sensory input. Psi researchers first sought to test this hypothesis by adapting the ganzfeld procedure, a mild sensory isolation technique introduced into experimental psychology during the 1930s.

In a ganzfeld telepathy experiment, one subject (the receiver) rests in a reclining chair in a soundproof chamber. Translucent Ping-Pong ball halves are taped over the eyes, and headphones are placed over the ears. A red floodlight is directed toward the receiver's eyes, and white noise is played through the headphones. (White noise is a random mixture of sound frequencies similar to the hiss made by a radio tuned between stations.) This homogeneous visual and auditory environment is called the Ganzfeld, a German

domly selected picture or videotaped sequence. For about 30 minutes, the receiver thinks aloud, providing a continuous report of all the thoughts, feelings, and images that pass through his or her mind. At the end of the ganzfeld period, the receiver is presented with several stimuli (usually four) and, without knowing which one was the target, is asked to rate the degree to which each matches the thoughts and images experienced during the ganzfeld period. If the receiver assigns the highest rating to the target, it is scored as a "hit." Thus, if the experiment uses judging sets containing four stimuli

Correcting the flaws

If the most frequent criticism of parapsychology is that it has not produced a repeatable psi effect, the second most frequent criticism is that many, if not most, psi experiments have inadequate controls and safeguards. A frequent charge is that positive results emerge primarily from initial, poorly controlled studies and then vanish as better controls and safeguards are introduced.

The potentially most fatal flaws in a psi study are those that would allow a receiver to obtain the target information in normal sensory fashion, either inadvertently or through deliberate cheating. This is called the problem of sensory leakage. Critic Hyman and parapsychologist Honorton agreed that the studies with good safeguards against sensory leakage obtained re-

Resolving disagreements over the existence of a disputed phenomenon has never been a matter of simply gathering more evidence.

(the target and three control stimuli), the hit rate expected by chance is one out of four, or 25 percent.

In 1985 and 1986, the *Journal of Parapsychology* devoted two entire issues to a critical examination of the ganzfeld studies, featuring a debate between Ray Hyman, a cognitive psychologist and a knowledgeable, skeptical critic of parapsychological research, and the late Charles Honorton, a prominent parapsychologist and major

sults that were just as strong as studies that had less good safeguards.

But because Hyman and Honorton disagreed on other aspects of the studies, they issued a joint communiqué in 1986, in which they agreed that the final verdict awaited the outcome of future experiments conducted by a broader range of investigators and according to more stringent standards. They then spelled out in detail the more stringent methodological and sta-

tistic
gove
ment
In
had
stud
guid
lishe
are
caus
perir
rand
the t
ceive
publ
of P
comj
was
of th
Janu
cal l
logic
T.
the r
catec
samo
stud
ings
it ha
or ar
psi
teste
dent
York

tistical standards they believed should govern all future ganzfeld experiments.

In 1983, Honorton and colleagues had initiated a new set of 11 ganzfeld studies that complied with all the guidelines he and Hyman later published in their joint communiqué. They are called autoganzfeld studies because a computer controlled the experimental procedures, including the random selection and presentation of the targets and the recording of the receiver's ratings. These studies were published by Honorton in the *Journal of Parapsychology* in 1990, and the complete history of ganzfeld research was resummarized by Bem (the author of this article) and Honorton in the January 1994 issue of the *Psychological Bulletin* of the American Psychological Association.

The autoganzfeld studies confirmed the results of the earlier, less sophisticated studies, obtaining virtually the same hit rate: about 35 percent. These studies also reconfirmed several findings from other research. For example, it has often been reported that creative or artistically gifted persons show high psi ability. The autoganzfeld studies tested this by recruiting twenty students from the Juilliard School in New York City to serve as receivers. Over-

all, these students achieved a hit rate of 50 percent, one of the highest rates ever reported for a single sample in a ganzfeld study. The autoganzfeld studies also found that significantly higher hit rates were obtained when the targets were videotaped film sequences rather than still pictures.

Belief and skepticism

Even if skeptical critics can agree that the autoganzfeld studies satisfy the strict methodological and statistical criteria set forth in Hyman and Honorton's joint communiqué, the studies cannot, by themselves, satisfy the further requirement that ganzfeld experiments be conducted by a broader range of investigators. In that sense, then, the jury is still out; the verdict awaits the outcome of future experiments. This state of affairs is not likely to change soon.

The history of science demonstrates that resolving disagreements over the existence of a disputed phenomenon has never been a matter of simply gathering more evidence until it reaches some objective, a priori threshold of quality and quantity. The amount of evidence required to persuade any given scientist that a phe-

nomenon exists depends on his or her belief as to how likely it is that the phenomenon exists in the first place.

Most scientists require more and better evidence for an anomalous phenomenon—one unexplained by known physical and biological mechanisms—than for other phenomena. This is usually expressed by the dictum that "extraordinary claims require extraordinary evidence." But in any given instance, there is no agreement on or objective measure of what constitutes "extraordinary."

Moreover, scientists' diverse reactions to evidence in disputed areas of research are strongly determined by their attitudes toward many other issues, not all of them strictly scientific. For example, scientists differ in the kinds of intellectual risks they are willing to take. For many scientists, it is far more sinful to conclude that an effect exists when it does not than to conclude that an effect does not exist when it does. The choice of which kind of error is more tolerable is not a matter of good science versus bad science but a matter of taste.

As Croesus learned the hard way, evaluating "psychic" data is inherently a risky business.