A Critique of Schwartz et al.'s After-Death Communication Studies

Studies with mediums by Gary Schwartz and colleagues have been widely reported in the media as scientific proof of life after death. But their experiments did not employ blind judging, used an inappropriate control group, and had insufficient safeguards against sensory leakage.

RICHARD WISEMAN and CIARAN O'KEEFFE

Schwartz, Russek, Nelson, and Barentsen (2001) recently reported two studies in which mediums appeared to be able to produce accurate information about the deceased under conditions that the authors believed "eliminate the factors of fraud, error, and statistical coincidence." Their studies were widely reported in the media as scientific proof of life after death (e.g., Matthews 2001; Chapman 2001). This paper describes some of the methodological problems associated with the Schwartz et al. studies and outlines how these problems can be overcome in future research.

Schwartz et al.'s first experiment was funded and filmed by a major U.S. television network (Home Box Office—HBO)

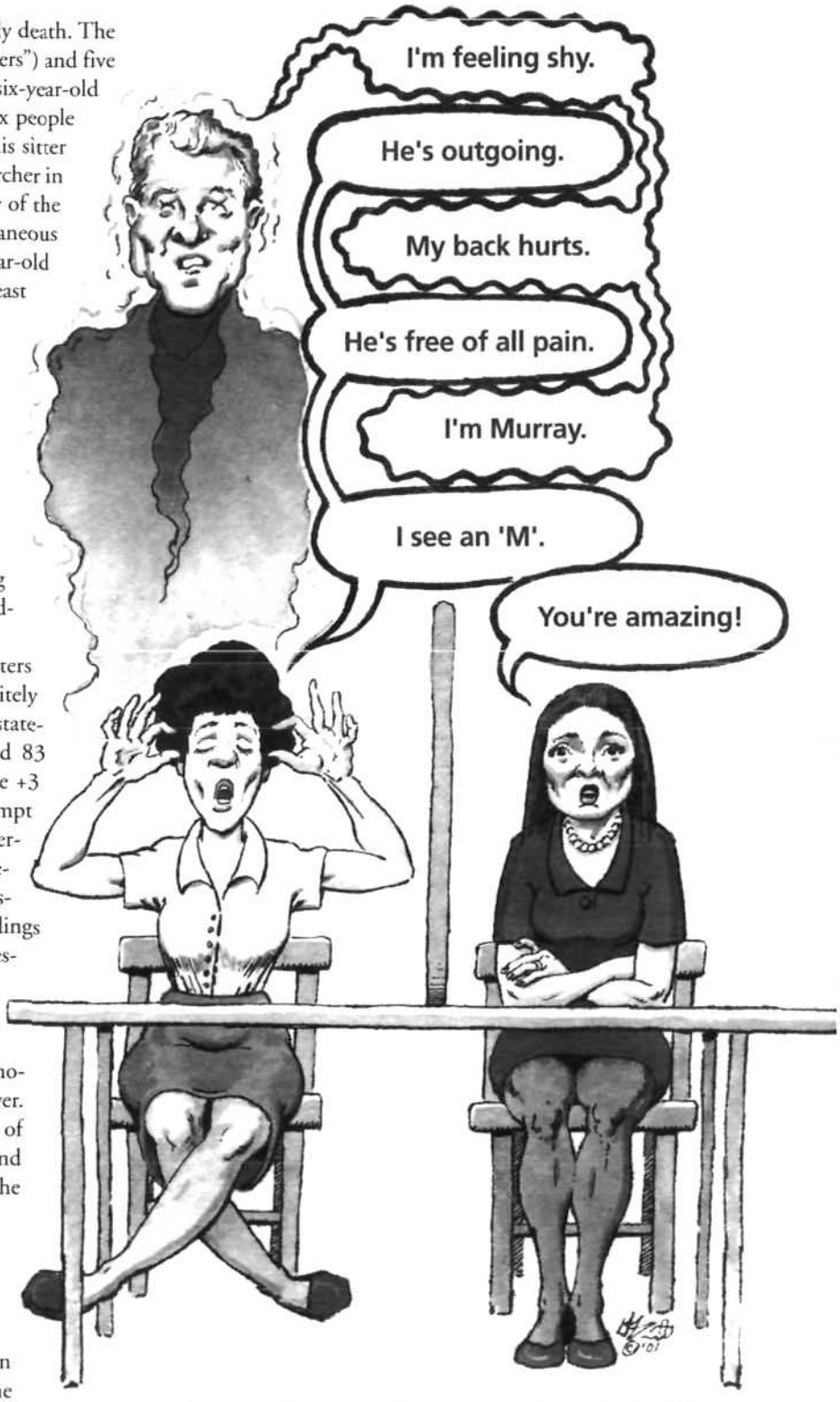
making a documentary about the survival of bodily death. The study involved two participants (referred to as "sitters") and five well-known mediums. The first sitter was a forty-six-year-old woman who had experienced the death of over six people in the last ten years. Schwartz et al. stated that this sitter was recommended to them by a well-known researcher in ADCs (After Death Communication) who "knew of the sitter's case through her research involving spontaneous ADCs." The second sitter was a fifty-four-year-old woman who had also experienced the death of at least six people in the last ten years.

During the experiment, the sitter and medium sat on either side of a large opaque screen. The medium was allowed to "conduct the reading in his or her own way, with the restriction that they could ask only questions requiring a yes or no answer." Each sitter was asked to listen to the reading and respond to the medium's questions by saying the word "yes" or "no" out loud. The first sitter was given a reading by all five mediums; the second sitter received readings from only two of the mediums.

A few months after the experiment, both sitters were asked to assign a number between -3 (definitely an error) to +3 (definitely correct) to each of the statements made by the mediums. The sitters placed 83 percent and 77 percent of the statements into the +3 category. Schwartz et al. also reported their attempt to discover whether "intelligent and motivated persons" could guess the type of information presented by the mediums by chance alone. The investigators selected seventy statements from the readings given to the first sitter and turned them into questions. For example, if the medium had said "your father loved dancing," the question became "Who loved to dance?" Sixty-eight undergraduates were shown these questions, along with a photograph of the sitter, and asked to guess the answer. Schwartz et al. reported that the average number of items guessed correctly was just 36 percent, and argue that the high level of accuracy obtained by the mediums could not be due to chance guessing.

The first sitter was then invited back to the laboratory to take part in a second experiment. In this experiment she received readings from two of the mediums who also participated in the first study. Rather than being separated by an opaque screen, the sitter sat six feet behind the medium. In the first part of these two readings the sitter was instructed to remain completely silent. In the second part she was asked to answer "yes" or "no" to each of the medium's questions. After reviewing the readings, the sitter rated 82 percent of the mediums' statements as being "definitely correct."

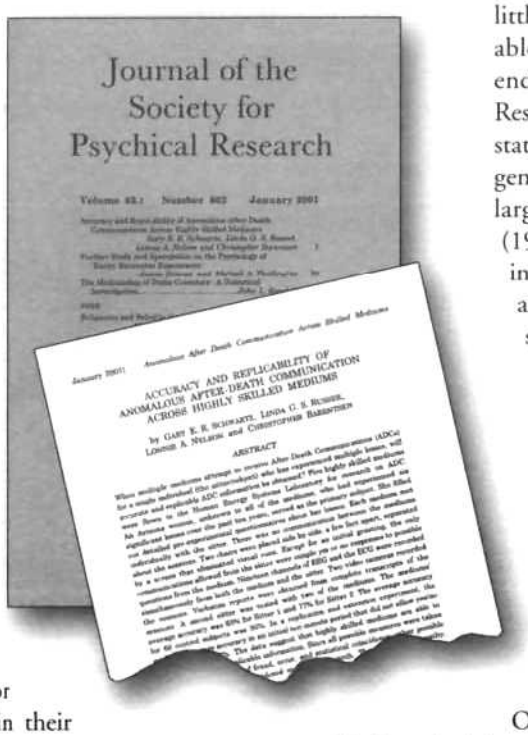
The Schwartz et al. studies suffered from severe methodological problems, namely: (1) the potential for judging bias, (2) the use of an inappropriate control group, and (3) inadequate safeguards against sensory leakage. Each of these problems will be discussed in turn.



Richard Wiseman and Ciaran O'Keeffe are in the Department of Psychology, University of Hertfordshire, College Lane, Hatfield Hert. AL10 9AB, U.K. E-mail: R.Wiseman@herts.ac.uk. This article also appears in the Paranormal Review.

Judging Bias

During a mediumistic reading the medium usually produces a large number of statements and the sitter has to decide whether these statements accurately describe his/her deceased friends or relatives. It is widely recognized that the sitter's endorsement of such statements cannot be taken as evidence of mediumistic ability, as seemingly accurate readings can be created by a set of psychological stratagems collectively referred to as "cold reading" (Hyman 1977; Rowland 1998). It is therefore vital that any investigation into the possible existence of mediumistic ability controls for the potential effect of these stratagems. Unfortunately, the Schwartz et al. study did not contain such controls, and thus it is possible that the seemingly impressive results could have been due to cold reading.



Schwartz et al. reproduced a small part of one reading in their paper, and this transcript can be used to illustrate how cold reading could account for the outcome of the studies. In the first line of the transcript the medium said, "Now, I don't know if they [the spirits] mean this by age or by generation, but they talk about the younger male that has passed. Does that make sense to you?" The sitter answered "yes." The medium's statement is ambiguous and open to several different interpretations. When the medium mentioned the word "younger" he/she could be talking about a young child, a young man, or even someone who died young (e.g., in their forties). The sitters may be motivated to interpret such statements in such a way as to maximize the degree of correspondence with their deceased friends and relatives if, for example, they had a strong belief in the afterlife, a need to believe that loved ones have survived bodily death, or were eager to please the mediums, investigators, and the HBO film crew.

In addition, the sitters may have endorsed the readings because some statements caused them to selectively remember certain events in their lives. As a hypothetical example, let us imagine that the medium had said, "Your son was an extrovert." This statement may have caused the sitter to selectively recall certain life events (i.e., the times that her son went to parties and was very outgoing), forget other events (e.g., the times that he sat alone and didn't want to be with others), and thus assign a spuriously high accuracy rating to the statement.

Biased interpretation of ambiguous statements and selective remembering can lead to sitters endorsing contradictory statements during a reading. Interestingly, the short transcript reproduced by Schwartz et al. contains an example of exactly this happening:

> Medium: . . . your dad speaks about the loss of child. That makes sense?
> Sitter: Yes.
> Medium: Twice? 'Cause your father says twice.
> Sitter: Yes.
> Medium: Wait a minute, now he says thrice. He's saying three times. Does that make sense?
> Sitter: That's correct.

Some of the statements made by the mediums may also have been true of a great many people and thus had a high likelihood of being endorsed by the sitters. For example, in the transcript the medium stated that one of the spirits was a family member, and elsewhere Schwartz et al. stated that the mediums referred to "a little dog playing ball." It is highly probable that many sitters would have endorsed both of these statements. Research has also revealed that many statements that do not appear especially general can also be true of a surprisingly large number of people. Blackmore (1994) carried out a large-scale survey in which more than 6,000 people were asked to state whether quite specific statements were true of them. More than one third of people endorsed the statement, "I have a scar on my left knee" and more than a quarter answered yes to the statement "Someone in my family is called Jack." In short, the mediums in the Schwartz et al. study may have been accurate, in part, because they simply produced statements that would have been endorsed by many sitters.

Other factors may also increase the likelihood of the sitter endorsing the mediums' statements. Clearly, the more deceased people known to the sitter, the greater chance they will have of being able to find a match for the medium's comments. Both sitters knew a relatively large number of deceased people. Both of them had experienced the death of six loved ones in the last ten years, and the first sitter reported that she believed that the mediums had contacted an additional nine of her deceased friends and relatives. Thus, the sitters' high levels of endorsement may have been due, in part, to them having a large number of deceased friends and relatives.

Control Group Biases

Schwartz et al. attempted to discover whether the seemingly high accuracy rate obtained by the mediums could have been the result of chance guesswork. However, the method developed by the investigators was inappropriate and fails to address the concerns outlined above. They selected seventy statements from the readings given to the first sitter in the first experiment and turned them into questions. For

example, if the medium had said "your son is very good with his hands," the question became "who was very good with his hands?" These questions were presented to a group of undergraduates, who were asked to guess the answers. Schwartz et al. reported that the average number of items guessed correctly was just 36 percent. However, it is extremely problematic to draw any conclusions from this result due to the huge differences in the tasks given to the mediums and control group. For example, when the medium said, "your son was very good with his hands," the sitter has to decide whether this statement matches the information that she knew about her deceased son. However, as noted above, this matching process may be biased by several factors, including her selective remembering and the biased interpretation of ambiguous statements. For example, the sitter may think back to the times that her son built model airplanes, endorse the statement, and the medium would receive a "hit." However, the control group were presented with a completely different task. They were presented with the question "Who was good with his hands?" and would only receive a "hit" if they guessed that the answer was the sitter's son. They therefore had a significantly reduced likelihood of obtaining a hit than the mediums.

Conceptually, this is equivalent to testing archery skills by having someone fire an arrow, drawing a target around wherever it lands and calling it a bullseye, and then testing a "control" group of other archers by asking them to hit the same bullseye. Clearly, the control group would not perform as well as the first archer, but the difference in performance would reflect the fact that they were presented with very different tasks, rather than a difference in their archery skills.

Psychical researchers have developed various methods to overcome the problems associated with "cold reading" when investigating claims of mediumistic ability (see Schouten 1994 for an overview). Most of these methods involve the concept of "blind judging." In a typical experiment, a small number of sitters receive a reading from a medium. The sitters are then asked to evaluate both his or her own reading (often referred to as the "target" reading) and the readings made for other sitters (referred to as "decoy" readings). If the medium is accurate then the ratings assigned to the target readings will be significantly greater than those assigned to the decoy readings. However, it is absolutely vital that the readings are judged "blind"—the sitters *should be unaware of whether they are evaluating a "target"* or "decoy" reading. This simple safeguard helps overcome all of the problems outlined above. Let us suppose that the medium is not in contact with the spirit world, but instead tends to use cold reading to produce seemingly accurate statements. These techniques will cause the sitters to endorse both the target and decoy readings, and thus produce no evidence for mediumistic ability. If, however, the medium is actually able to communicate with the spirits, the sitters should assign a higher rating to their "target" reading than the "decoy" readings, thus providing evidence of mediumistic ability.

It is hoped that future tests of mediumistic ability will employ the type of blind judging methods that have been developed, and frequently employed, in past tests of mediumistic ability.

However, blind judging is only one of several methodological *safeguards that should be employed when testing mediumistic* ability. Well-controlled tests should also obviously prevent the medium from being able to receive information about a sitter through any normal channels of communication. Unfortunately, the measures taken by Schwartz et al. to guard against various forms of potential sensory leakage appear insufficient.

Mediums in the Schwartz et al. study may have been accurate, in part, because they simply produced statements that would have been endorsed by many sitters.

Sensory Leakage

Throughout all of the readings in the first experiment, and the latter part of the readings in the second experiment, the sitter *was allowed to answer "yes" or "no" to the medium's questions.* These answers would have provided the mediums with two types of information that may have helped them produce more accurate statements in the remainder of the reading. First, it is very likely that the sitter's voice would have given away clues about her gender, age, and socioeconomic group. This information could cause the mediums to produce statements that have a greater likelihood of being endorsed by the sitter. For example, an older sitter is more likely to have experienced the death of their parents than a younger sitter, and certain life events are gender-specific (e.g., being pregnant, having a miscarriage, etc.). Second, the sitters' answers may have also given away other useful clues to the mediums. For example, let us imagine that the medium stated, "I am getting the impression of someone male, is that correct?" If the sitter has recently lost someone very close to her, such as a father or son, then she might answer a tearful "yes." If, however, the deceased male was an uncle that sitter didn't really know very well, then her "yes" might be far less emotional. Again, a skilled medium might be able to unconsciously use this information to produce accurate statements *later in the reading. Any well-controlled test of mediumistic* ability should not allow for the sitter to provide verbal feedback to the medium during the reading.

In the first part of the readings in the second experiment, the sitter was asked not to answer yes or no to any of the medium's statements. However, the experimental set-up still employed insufficient safeguards against potential sensory leakage. The medium sat facing a video camera and the sitter sat six feet behind the medium without any form of screen separating the two of them. As such, the sitter may have

emitted various types of sensory signals, such as cues from her movement, breathing, odor, etc. Parapsychologists have developed elaborate procedures for eliminating potential sensory leakage between participants (e.g., Milton and Wiseman 1997). These safeguards frequently involve placing participants in separate rooms, and often the use of specially constructed sound-attenuated cubicles. Schwartz et al. appeared to have ignored these guidelines and instead allowed the sitter to interact with the medium, and/or simply seated them behind one another in the same room. Neither of these measures represent sufficient safeguards against the potential for sensory leakage.

The investigators also failed to rule out the potential for sensory leakage between the experimenters and mediums. The second sitter in the first experiment is described as being "personally known" to two of the experimenters (Schwartz and Russek). The report also described how, during the experiment, the mediums were allowed to chat with Russek in a courtyard behind the laboratory. Research into the possible existence of mediumistic ability should not allow anyone who knows the sitter to come into contact with the medium. Schwartz allowed such contact, with their only safeguard being that the mediums and Russek were not allowed to talk about matters related to the session. However, a large body of research has shown that there are many ways in which information can be unwittingly communicated, via both verbal and nonverbal means (e.g., Rosenthal and Rubin 1978). As such, the safeguards employed by Schwartz et al. against possible sensory leakage between experimenter and mediums were insufficient.

In short, the Schwartz et al. study did not employ blind judging, employed an inappropriate control group, and had insufficient safeguards against sensory leakage. As such, it is impossible to know the degree to which their findings represent evidence for mediumistic ability. Psychical researchers have worked hard to develop robust methods for testing mediums since the 1930s (see Schouten 1994). It is hoped that future work in this area will build upon the methodological guidelines that have been developed and thus minimize the type of problems discussed here.

References

Blackmore, S. 1994. Probability Misjudgement and belief in the paranormal: Is the theory all wrong? In D. Bierman (Ed.), *Proceedings of the 37th Annual Convention of the Parapsychological Association* 72–82.

Chapman, J. 2001. Is there anybody there? Mediums perform well in scientific séance test. *The Daily Mail*, March 5.

Hyman, R. 1977. Cold reading: How to convince strangers that you know all about them. SKEPTICAL INQUIRER 1(2): 18–37.

Matthews, R. 2001. Spiritualists' powers turn scientists into believers. *The Sunday Telegraph*, March 4.

Milton, J., and R. Wiseman. 1997. Guidelines for extrasensory perception research. University of Hertfordshire Press: Hatfield, England.

Rosenthal, R., and D.B. Rubin. 1978. Interpersonal expectancy effects: The first 345 studies. *Behavioural and Brain Sciences* 3, 377–386.

Rowland, I. 1998. *The Full Facts Book of Cold Reading*. Ian Roland Limited: London, England.

Schouten, S.A. 1994. An overview of quantitatively evaluated studies with mediums and psychics. *The Journal of the American Society for Psychical Research*, 88:221–254.

Schwartz, G.E.R., L.G.S. Russek, L.A. Nelson, and C. Barentsen. 2001. Accuracy and replicability of anomalous after-death communication across highly skilled mediums. *Journal of the Society for Psychical Research*, 65(862):1–25.