

THE AMERICAN INSTITUTES FOR RESEARCH REVIEW OF THE DEPARTMENT OF DEFENSE'S STAR GATE PROGRAM: A COMMENTARY

BY EDWIN C. MAY

ABSTRACT: As a result of a Congressionally Directed Activity, the Central Intelligence Agency (CIA) conducted an evaluation of a 24-year, government-sponsored program to investigate ESP and its potential use within the intelligence community. The American Institutes for Research (AIR) was contracted to conduct the review of both research and operations. Their September 29, 1995 final report was released to the public November 28, 1995. As a result of AIR's assessment, the CIA concluded that a statistically significant effect had been demonstrated in the laboratory but that there was no case in which ESP had provided data that had ever been used to guide intelligence operations. This paper is a critical review of AIR's methodology and conclusions. It will be shown that there is compelling evidence that the CIA set the outcome with regard to intelligence usage before the evaluation had begun. This was accomplished by limiting the research and operations data sets to exclude positive findings, by purposefully not interviewing historically significant participants, by ignoring previous extensive Department of Defense program reviews, and by using the questionable National Research Council's investigation of parapsychology as the starting point for their review. Although there may have been political and administrative justification for the CIA not to accept the government's in-house program for the operational use of anomalous cognition, these external considerations appeared to drive the outcome of the evaluation. As a result, they have come to the wrong conclusion with regard to the use of anomalous cognition in intelligence operations and have significantly underestimated the robustness of the basic phenomenon.

As part of the fiscal year 1995 defense appropriations bill, responsibility for the government-sponsored investigation and use of ESP¹ was transferred to the Central Intelligence Agency (CIA). In a Congressionally Directed Action, the CIA was instructed to conduct a retrospective review of the 24-year program, now known as Star Gate, that resided primarily within the intelligence community. The analysis was to include the research that was conducted since 1972 at SRI International and later at Science Applications International Corporation (SAIC). In addition, the CIA was to include an assessment of the intelligence-gathering utility of anomalous cognition (AC), and the program history was to be

¹ Hereafter I use the term *anomalous cognition* (AC) instead of ESP. Anomalous cognition is defined as a form of information transfer in which all known sensorial stimuli are absent. This is also known as remote viewing (RV) and clairvoyance.

declassified. Initiated in June 1995, the evaluation was to be completed by September 30, 1995.

The CIA contracted with the American Institutes for Research (AIR) to manage the review. They, in turn, formed a "blue-ribbon" panel that included psychologist Ray Hyman from the University of Oregon and statistician Jessica Utts from the University of California at Davis. AIR contributed Michael Mumford and Andrew Rose to the panel to provide unbiased assessment on methodological issues. The president of AIR, David Goslin, served as coordinator of the evaluation effort.

I was asked by the CIA to provide administrative support, technical documents, and briefings on an as-needed basis for the review. This work was supported by a small contract to SAIC.²

The CIA-sponsored AIR investigation concluded that a statistically significant laboratory effect had been demonstrated but that more replications were needed. In no case had the AC information ever been used to guide intelligence operations (Mumford, Rose, & Goslin, 1995).

I question the validity of their and the CIA's conclusions because they:

1. Limited the data sets in the analysis. As a way of officially ignoring AC's positive contributions to intelligence, only a small fraction of the operational remote-viewing database was examined. That was the final data collected just before the unit closed, a time widely known as problematic. In their laboratory evaluations, they restricted the investigation to only the government-sponsored research and then insisted on the need for more outside replications. In doing so, they ignored the conclusions of one of their own investigators who showed that the government-sponsored research had already been conceptually replicated.

2. Failed to contact significant program participants. Because of the complexity of the 24-year program, it is impossible to conduct an in-depth and accurate evaluation without significant contact with the program's many major participants. The review focused on the project's reports, but they were written to satisfy specific contract requirements and were not designed individually or in total to serve as a program justification; thus, these documents provide a substantially incomplete picture of the program.

In addition to questioning the validity of CIA/AIR's conclusions, I find such serious problems with their evaluation methodology that I have become reluctantly convinced that their conclusions were set before their investigation began. The investigators failed to:

² I resigned from SAIC on November 28, 1995, and do not speak for SAIC or for any of their clients.

1. Apply consistent criteria for acceptance or rejection of AC. The investigators were troubled by possible non-AC alternative explanations for the statistically significant laboratory results, yet they ignored similar alternatives for the failed operations. For example, well-known psychological effects such as bad morale, failed expectations, and a lack of a supportive environment were not discussed as potential alternatives for the failed operations. In their positive forms, all of these psychological effects are critical for excellence in any human activity.

2. Avail themselves of the previous exhaustive reviews conducted by various organizations within the DOD, all but one of which was positive. Since the CIA was allowed only 4 months to complete the evaluation, it is surprising that they chose not to use this resource.

3. Reject a discredited evaluation of parapsychology conducted by the National Research Council (NRC). They knew that the NRC investigators were not cleared for access to the vast majority of SRI's research, yet the AIR investigation relied heavily on the NRC's review to question the SRI research results prior to 1988.

4. Use neutral government scientific evaluation resources such as the Military Services' or the CIA's Scientific Advisory Boards. Instead they commissioned external investigators with previously published conclusions about parapsychology. The CIA could then justify whatever conclusion they wished, because it would be consistent, by definition, with at least one of their external reviewers.

5. To recognize a potential significant conflict of interest for David Goslin, president of AIR and a report co-author. He had administrative responsibility for the discredited NRC investigation of parapsychology.

Finally, since the political situation and the status of the program had significantly deteriorated technically and administratively, I speculate that this contributed to the underlying reason why the CIA did not want the program even before the evaluation began.

In this paper, I will expand upon these topics to demonstrate clearly that the outcome and conclusions drawn by AIR and subsequently the CIA were set before the investigation began, and that methodological and administrative choices were made to assure that the results of the investigation would support the CIA's predetermined perspective. In addition, I will document that they have come to the wrong conclusion with regard to the use of AC in intelligence operations and greatly underestimated the robustness of the phenomenon.

CRITIQUE OF THE CIA/AIR CONCLUSIONS

Limited Database for the Evaluation of Research and Operations

The program evaluation was set from the beginning to include only government-sponsored research. If the evaluation had been confined to the assessment of the scientific quality of the research, then perhaps this was not a bad idea, given that the Congress was trying to determine whether there was merit to continue. Upon closer inspection, however, even in this case, limiting the scope of the evaluation to exclude replications is scientifically invalid. The evidence for or against a statistically based phenomenon cannot rest on the evidence provided by a few investigators in two laboratories (i.e., SRI and SAIC). Rather, science demands that the evidence rest in replications. Yet the reviewers were requested not to look outside the Star Gate project.

In the CIA's briefing to Congress, they list three points as attributed to the AIR investigation³

1. "The data do not establish that a paranormal phenomenon is involved, nature of source not identified."
2. "The data have not been replicated independently."
3. "The boundary constraints critical to obtaining statistically significant experimental results are not practical in real world of intelligence collection."

No statistically based phenomena can be established without replication, yet the investigators were instructed not to look for any. (Utts ignored this instruction and clearly showed that a conceptual replication has been well established in the literature and that significant statistical consistencies existed between the SRI and SAIC data sets.) Since the investigators were restricted at the outset, the top two items above are true by construction, not by analysis.

A casual scan of my collection of technical journals found four independent replications of remote viewing (Dunne & Bisaha, 1979; Schlitz & Gruber, 1980; Schlitz & Haight, 1984; and Targ, Katra, Brown, & Wiegand, 1995). Rather than more replications as called for by AIR and Hyman, what is needed is a meta-analysis of all the AC studies to date and more attention on potential mechanisms.

Perhaps I should rest my case here. The CIA/AIR conclusions appeared to be designed into the investigation. Their final item above is questionable on its face value because it is true by the nature of

³ I quote the following points from a personal communication to me by a U. S. Senate Appropriations Committee staff member.

intelligence, not because of a valid criticism of the program's operational AC. The only valid measure of intelligence utility for AC is a top-level outcome measure, not a statistical analysis. In short, do end-users come back for more? Do any end-users have cases that they can point to that helped solve an intelligence problem? The CIA and AIR say no, but as I will show, that conclusion was also arrived at by construction rather than by analysis.

I first learned of the CIA/AIR's plan for the evaluation of the intelligence value of AC from Mumford during the July meeting of the blue-ribbon panel at which I was invited to present material and answer questions. At that date, Mumford claimed that they were only going to look back 3 years from the end of the 24-year program. I told him that I was convinced that this would not provide an honest picture of the utility of AC. I informed the panel that I could easily predict the outcome from my knowledge of the morale of the government's viewers, the standard management by Defense Intelligence Agency (DIA) officials, the tasking (i.e., what data they were after), and the inappropriate collection protocols.

Mumford attempted to justify his decision by saying he did not want to rely on memory and hearsay. He would not have to, because there was an extensive written history including testimonials to official organizations within the intelligence community. Mumford reiterated that he was sticking to his plan, regardless.

I objected to this decision to ignore existing data. I called the individual at CIA who had been assigned to manage the review, hereafter called the Point of Contact or POC, and insisted that some of the officials I had previously named had to be contacted. I learned later that the names and phone numbers of at least six individuals had been given to the POC. These end-users, some on active duty and some retired, have already been on written record as attesting to the value of AC-derived intelligence data in solving specific problems.

After the AIR report had been given to Congress, but before it was released to the public and before I had seen it, I called many of the individuals on the list. Most had not been contacted, and those who had been told the CIA representative the case specifics and value of their individual intelligence application examples. Some of the positive findings occurred before the final year but within the last 3 years of the project.

Finally, even a cursory investigation of the written record of intelligence operations would have revealed substantial evidence of the operational utility of AC. Minimally, there exist enough data to claim *prima facie* utility with regard to the method, and selected cases are beyond doubt as to AC's specific utility.

Joseph McMoneagle, one of the original government viewers beginning in 1978 and a consultant to the SRI/SAIC and Cognitive Sciences Laboratory, in 1984 was granted a Legion of Merit award for excellence in intelligence service. The Legion of Merit is rarely awarded for other than 20 or 30 years' service, yet McMoneagle received his on the following basis. I quote, with permission, from McMoneagle's citation:

He [McMoneagle] served most recently as a Special Project Intelligence Officer for SSPD, SSD, and 902d MI Group, as one of the original planners and movers of a unique intelligence project that is revolutionizing the intelligence community. While with SSPD, he used his talents and expertise in the execution of more than 200 missions, addressing over 150 essential elements of information [EEI]. These EEI contained critical intelligence reported at the highest echelons of our military and government, including such national level agencies as the Joint Chiefs of Staff, DIA, NSA, CIA, DEA, and the Secret Service, producing crucial and vital intelligence unavailable from any other source.

How is it that the CIA and AIR could not find compelling evidence for the operational utility of anomalous cognition? They clearly chose not to look.

Important Program Participants: Never Contacted

From 1985 through 1990, the research project at SRI International enjoyed substantial, on-going, and written scientific oversight of the major portion of the AC database at SRI. Twelve individuals who are world-renowned in their individual disciplines were chosen by the client and other government officials to serve on our Scientific Oversight Committee. They were selected on the basis of their scientific reputations and on the basis of their skepticism. "Believers" were not allowed on the committee. The SOC's responsibilities were fourfold:

1. to evaluate our written protocols prior to our conducting any experiments. The protocol that was actually used for each investigation was the convergence of round-robin exchange with the SOC.
2. to exercise unannounced drop-in privileges to observe experiments in progress. Approximately one half of the SOC availed themselves of this opportunity.
3. to review the then classified final research reports as if the reports were technical journal submissions in the individual disciplines of the committee members. The disciplines included physics, philosophy, psychology, electrical engineering, statistics, and astronomy. Their reviews were in writing and appended, unedited, to each of our final reports.

4. to suggest approaches for research in the next year of the 5-year contract.

During the SAIC time, the SOC was limited to only five members, but they had the same charter. Three of the five came from the SOC at SRI. At SAIC we established two additional oversight committees. An Institutional Review Board (i.e., human use committee) was established with nine members who were health and medical professionals and are renowned in their disciplines as well. This committee included one Nobel laureate (as did SAIC's SOC). Besides assuring the protection of our human subjects, they also served as a less formal scientific oversight committee.

The third oversight committee at SAIC was for policy. The three members of this committee came from formerly very senior positions in the DOD, and their job was to assure that we were meeting our obligations to the DOD and supporting its mission.

Of these 17 individuals who had intimate knowledge of the inner workings of this project, scientifically, methodologically, and administratively, only one was contacted by the CIA. It was that single individual who provided the names of satisfied end-users that I discussed above.

The SOC's comments were available to the AIR reviewers in written form. Many of the committee members lived on the east coast, and a few even lived in Washington. The CIA/AIR investigators could have easily contacted them. They didn't.

The failure to contact significant program participants does not end with these committees. I provided the POC with the names and phone numbers of numerous other pertinent individuals. The list included the previous project director for Star Gate who had retired less than a year before the review, and the Commander for a still-classified client who initiated a single contract that accounted for a significant fraction of all the funding for the project over the 24 years. In addition, I gave the POC the names of a number of the original government viewers. In short, with interviews of mostly local people, the CIA could have gained significant insight into the scientific, operational, managerial, and political aspects of the Star Gate project. They chose to ignore these resources.

METHODOLOGICAL PROBLEMS

Inconsistent Criteria

One of AIR's significant methodological flaws is important with regard to the assessment of operations. In the section on the Evaluation

Plan in the report, Mumford et al. (1995, p. 2-1) correctly required of the laboratory investigations "*unambiguous* [emphasis added] evidence for the existence of the phenomenon." Following this lead, Hyman hypothesized a number of alternative explanations for the observed statistical significance other than the AC one, although he admits he couldn't find any obvious flaws in the methodology (Mumford et al., 1995, p. 3-75). For example, he is troubled that during the SAIC research, a single judge was used to conduct all the laboratory evaluations. Although Hyman does not propose how this might affect the result, he is correct in proposing the hypothesis that it might somehow effect the outcome. (Hyman lists other alternatives as well, but this one illustrates the point.) As it turns out, Utts finds statistical homogeneity (i.e., meaningful consistency) among the results from SRI, SAIC, and replications elsewhere when that single judge was not involved. Thus, this hypothesis must be rejected. This same consistency also rejects the other alternatives Hyman proposes as well.

Yet, the AIR failed to apply the same "unambiguous" criteria to their evaluation of the efficacy of AC in intelligence operations—in this case, why operations may have failed. In particular, in their discussion in the section "Evaluating the Utility of Remote Viewing in Intelligence Operations," they list a number of "boundary conditions" that might affect AC in operations. These include a number of physical and methodological issues such as feedback and whether a sender or distance to the target might be factors.

They did not discuss or propose any psychological issues that may have been the deciding factors as to why the operations failed in their limited sample. For example, it is well known that human performance of any kind and most certainly AC performance is profoundly affected by the morale and expectations of the participants and by the emotional environment in which the performance is expected (e.g., home-team effect in sports). But none of these potentially critical factors was discussed in the context of reaching the unambiguous conclusion that AC was useless in operations.

I had discussed these points in my meeting with the blue-ribbon panel in July 1995. In particular, having spent considerable time with the government remote-viewing unit, I was knowledgeable about what psychologists call "set and setting." That is, I saw first hand and reported to the panel that during the last 2 years (i.e., the time of the operational evaluation) the emotional environment had deteriorated to the point that the viewers wanted to leave the unit, and some of the staff had already left in disgust.⁴ The morale was so low that doing excellent remote

⁴ This was reported to me by the foreign analyst for the unit in a personal communication in June 1995.

viewing, or practically anything else, would be out of the question. The AIR investigators interviewed the government remote viewers (Mumford et al., 1995, p. 4-9) and learned of these problems first hand.⁵

These critically important factors were completely left out of the discussion in the report and no alternate hypotheses were suggested to question their unambiguously negative conclusion about the use of AC in intelligence operations.

Previous Program Reviews

Even before I was officially under contract with the CIA, I gave the POC either copies of, or pointers to, a number of classified program reviews that had been conducted in the past.⁶

One important aspect of the program was its ongoing and rigorous review and technical oversight. Everyone involved (i.e., the government sponsors, SRI, and SAIC) were correctly concerned that the research should be as rigorous as possible and that the program could be justified within the intelligence community and DOD. These reviews were extensive and were conducted by general military officers, senior members of the intelligence community, respected scientists from many disciplines, and end-users of the AC intelligence product.

These reviews remain classified and, with one exception, were positive with regard to the existence of AC and its successful contributions to intelligence. Even the negative one wanted only to stop the research but continue the operations! The final such review was conducted in 1987.

In addition to the written reviews, from 1985 through 1990 the program enjoyed the continued oversight of a high-ranking military officer from the still-classified sponsor and a GS-15 geneticist from DIA as permanent on-site observers at SRI.

The POC is a PhD scientist who at the time seemed dedicated to the best job possible. He informed me, however, that the CIA intended to ignore the previous reviews and start fresh. Given that the review had to be in Congress in 4 months, I could not conceive how it could be effective and accurate and ignore the substantial amount of previous oversight. After all, a complete analysis could, and should, have included a review of the previous classified DOD assessments.

⁵ This was reported to me by one of the government remote viewers in a personal communication in August 1995.

⁶ There have been a number of unclassified reviews and published meta-analyses of anomalous mental phenomena, but they do not deal with operations.

A Thread of Bias, Potential Conflict of Interest, and Suppression of Data

In the early days of the project, Targ and Puthoff (1974a) reported on a series of experiments they conducted at SRI with Uri Geller, an Israeli magician/psychic. George Lawrence from the Advanced Research Projects Agency (ARPA) accompanied by two consultants, Ray Hyman and Robert Van de Castle, came to SRI requesting an opportunity to see an experiment in progress with Geller. Puthoff and Targ correctly denied access to the ARPA representatives because of technical and administrative protocol issues. After all, with such controversy swirling about Geller, it was easy to become quite paranoid about who was trying to trick whom. The safest and the most scientifically sound course was not to allow anyone except the direct research team to witness formal experiments regardless of credentials (H. Puthoff, personal communication, 1996; Targ & Puthoff, 1977).

Yet, as part of their cover story, *Time* magazine (Jaroff, 1974) quoted Ray Hyman's claim that the SRI tests were carried out with "incredible sloppiness." The irony is that the tests that Hyman and Lawrence witnessed at SRI were indeed conducted with incredible sloppiness, but the experiments they witnessed were of their own making and had nothing at all to do with protocols of those experiments to which they had been denied access (H. Puthoff, personal communication, 1996; Targ & Puthoff, 1974b). It is clear that Lawrence and Hyman had strongly held positions and were willing to report their experiences at SRI inaccurately. Thus we see the first evidence of a negative bias on the part of Lawrence and Hyman.

In 1984 their biases were again demonstrated. The Army Research Institute (ARI) commissioned the National Academy of Sciences to investigate the potential of certain techniques that propose to enhance human performance (Druckman & Swets, 1988). Although performance enhancement has never been the claim of research parapsychology, the National Research Council (NRC) included parapsychology as one of the topics to be studied. The same George Lawrence, formerly from ARPA, was ARI's project monitor, and he asked that Ray Hyman be commissioned to head the investigation into parapsychological phenomena. David Goslin, Executive Director of the Commission on Behavioral and Social Sciences and Education for the NRC, served as overall project director and agreed to the request.

Regarding parapsychology, the NRC study concluded (Druckman & Swets, 1988):

The committee finds no scientific justification from research conducted over a period of 130 years for the existence of parapsychological

phenomena. It therefore concluded that there is no reason for direct involvement by the Army at this time. We do recommend, however, that research in certain areas be monitored, including work by the Soviets and the best work in the United States. The latter include that being done at Princeton University by Robert Jahn; at Maimonides Medical Center in Brooklyn by Charles Honorton, now in Princeton; at San Antonio by Helmut Schmidt; and at the Stanford Research Institute by Edward [sic] May. Monitoring could be enhanced by site visits and by expert advice from both proponents and skeptics. The research areas included would be psychokinesis with random event generators and Ganzfeld effects. (p. 22)

By the time the NRC began their investigation, I was the project director at SRI International. Our program was highly classified at that time, and special access was required before any aspect of the project could be discussed even with individuals with appropriate security clearances. Thus, neither the in-house DIA classified program nor the NRC investigators, and particularly Ray Hyman, had access to over 80% of all the remote viewings conducted during the SRI years. None of the research reports from this contract were kept with the DIA remote-viewing group. So even though Hyman had access to this group, he was denied access to, and was probably even unaware of, the SRI data of that time period.⁷

I was not even allowed to meet with Hyman in our laboratory or office space; he and I met in a separate building at SRI that was not associated with the project. Our discussions were confined to our published account of a careful random number generator experiment that we had conducted in 1979.⁸

In the summary conclusion of the NRC quoted above, remote viewing was not even mentioned, although an analysis of the early studies at SRI and later studies at Princeton is contained in the body of the NRC report. Their conclusion on remote viewing was that "the literature on remote viewing has managed to produce only one possibly successful experiment that is not seriously flawed in its methodology—and that one experiment provides only marginal evidence for the existence of ESP" (p. 22).

The parapsychology section of the NRC study was a mockery of good science and serves as an excellent model for a pseudoscientific investigation. The methodology for the NRC investigation and the NRC conclusions were soundly criticized and shown to be without scientific merit

⁷ The research from 1985 through 1990 was finally declassified as a result of my petition in 1990. Many of the sponsors and most of the intelligence operations, however, remain classified.

⁸ It is curious to note that our RNG study was singled out in the NRC report as "singularly well controlled" and "the only near-flawless RNG experiment" (Druckman & Swets, 1988, p. 189).

(Palmer, Honorton, & Utts, 1989). Palmer et al. (1989) summarized their four major points:

1. The scope of the review is limited to less than 10% of the systematic scientific effort in parapsychology
2. The two principal evaluators of parapsychological research for the Committee, Ray Hyman and James Alcock, were publicly committed to a negative position⁹ on parapsychology at the time the Committee was formed
3. . . . the Committee's method of assessing parapsychology violates its own stated guidelines for research evaluation, which specify the identification and assessment of plausible alternatives. With regard to the better parapsychological experiments, the Committee admits, "We do not have a smoking gun, nor have we demonstrated a plausible alternative" ([Druckman & Swets, 1988,] p. 200).
4. The report selectively omits important findings favorable to parapsychology contained in one of the background papers commissioned for the Committee, while liberally citing from other papers supportive of the Committee's [negative] position. The principal author of the favorable paper, an eminent Harvard psychologist, was actually asked by the Chair of the NRC Committee to withdraw his favorable conclusions. (p. 32)

This last point is particularly troublesome and reveals the political nature of what should have been a carefully conducted scholarly investigation of the kind that usually characterizes the NRC. Violating one of the basic tenets of science to report all findings, the NRC Committee asked Robert Rosenthal to

omit the section of our paper evaluating the Ganzfeld research domains. I refused to do so but was so shocked and disappointed by this request that I discussed this request with a number of colleagues in the Harvard departments of Psychology and of Statistics. Without exception they were as shocked as I was.

In the end, censorship did not occur, and Monica Harris' and my paper is available in its entirety in a kind of preprint format from the National Academy Press. (Quoted with permission from a letter from Professor Rosenthal to me, November 28, 1995)

Rosenthal's and Harris's commissioned paper listed the Ganzfeld methodological quality to be superior to the typical quality of the other four areas they considered (Rosenthal, 1990).

In addition to the significant methodological flaws and the attempt to suppress positive findings, the NRC study was essentially contradicted in

⁹ In addition, the sentence "the total accumulation of 130 years' worth of psychical investigation has not produced any consistent evidence for paranormality" can be found in Hyman (1985, p. 7) and in the NRC conclusion (Druckman & Swets, 1988, p. 22) and thus demonstrates Hyman's stated bias before the NRC investigation was complete.

its major conclusion by a one-day workshop hosted by the Office of Technology Assessment, the research arm of the U.S. Congress (Office of Technology Assessment, 1989). The OTA did not completely exonerate the field of parapsychology; there is no scientific endeavor that cannot be improved. The OTA did, however, clearly demonstrate that the research cannot simply be dismissed—a view directly opposite to the NRC's conclusion.

In outlining the development of a potential conflict of interest, I point out once again that David Goslin had administrative responsibility for this seriously flawed NRC investigation.

When the CIA was searching for someone to conduct their technical review of the Star Gate program, they were turned down by the NRC in part because of the time constraint and in part because of the substantial negative publicity that resulted from their previous report on parapsychology.¹⁰ Instead, AIR was commissioned to conduct the review. AIR's president is David Goslin.

Let me now summarize the thread of bias and potential conflict of interest. Ray Hyman and George Lawrence were denied access to SRI experiments with Uri Geller in 1974. Ray Hyman has a long history of a negative bias with regard to parapsychology. In 1985 George Lawrence commissioned the National Academy of Sciences to investigate parapsychology and picked Hyman to direct the effort. In 1986 David Goslin presided over a methodologically flawed review. In 1995 David Goslin assumed responsibility for the CIA-sponsored investigation of the Star Gate program.

It is not surprising that the NRC study is liberally quoted in the AIR report because it supports the possibly predisposed views of CIA/AIR, albeit from a flawed investigation. Since Jessica Utts was one of the coauthors of the formal response to the NRC study, I questioned her (in an e-mail message to her of December 17, 1995):

Since you were a contributing author to the reply [to the NRC investigation] and since the reply soundly criticized the NRC's review methodology, I was surprised to see that you did not mention the NRC study or the PA's [Parapsychological Association] reply in your section of the AIR's report. Considering the weight that the AIR investigators placed on the NRC study, I feel it was a substantial oversight for you not have added your first-hand criticism of the NRC report as part of your remarks.

So that I make no errors in interpretation, I print, with permission, her complete reply to me of December 19, 1995:

¹⁰I learned this in a 1995 personal communication to me from the CIA point of contact.

This is in response to your question about why I did not mention the National Research Council's 1988 evaluation of parapsychology in my report to AIR. The answer is that I was explicitly asked by AIR staff NOT to mention the NRC report in my review! This is very troubling to me for a number of reasons.

First, you are correct in stating that I was aware that the NRC committee was not shown much of the relevant remote viewing data when they did their review, and that they did not in fact even know the data existed. As you also noted, I co-authored a critical review of the NRC report shortly after it was published, illustrating a number of weaknesses with it.

What you may not know is that in addition to those problems, the statistical method the NRC committee relied on for its findings (called "vote-counting") has been completely discredited, and is known to produce misleading results. I raised this point at the July meeting Ray Hyman and I attended with the AIR staff at their Palo Alto office, and it was substantiated by Stanford Statistics Professor Lincoln Moses, who had been asked by the AIR staff to attend the meeting to comment on that and related statistical issues. (Had the NRC committee included a statistician, that serious flaw, and the subsequent misleading results, may have been avoided. I am sorry to say that even at our meeting in Palo Alto, Ray did not seem to understand the problem, and he was the principal "statistician" for the NRC report.)

When I was explicitly asked by AIR staff NOT to mention the NRC report in my review, I assumed they had realized the problems with it, and, especially given the involvement of the AIR President with the NRC Committee, were happy to let it fade into oblivion.

Given that background, I was quite disappointed to see that AIR made liberal use of the NRC report in their conclusions. Had I known they were going to do that, I certainly would have discussed the multiple problems with it in my report. By not mentioning it, an uninformed reader may assume that I support it, which I certainly do not.

I would also like to explain another omission in my report that occurred for much the same reason. Despite the claims Ray Hyman is making in the media, we were shown very little of the "operational" remote-viewing work. One of the few documents we were shown was a list of [the former DIA project officer's] "best" remote viewing successes. Since the list provided almost no detail, you may recall that I asked you for names and numbers of individuals I could contact to get more information about those purported operational successes. In a memo dated August 1, 1995, you provided me with phone numbers for [a former DIA project officer, a former senior DIA official, a military General who had program responsibility], and Joseph McMoneagle. You sent a copy of the memo to the AIR staff.

Shortly after you sent me that memo, I was contacted by the AIR staff and told that I was NOT to contact any of those individuals. Thus, I was not able to gain any details about the operational remote viewing work. I thought you should know that, in case you were wondering why I requested that

information and then did not use it. Again, I am clueless as to why Ray Hyman is making claims in the media that we had access to the operational work for our review. I do not think he was given access to any information not shown to me. I don't know how he can substantiate the claims he's making about remote viewing being useless for intelligence. He may be correct, but he has very little data on which to base that conclusion.

Although a case can be made that Utts should not be contacting people with regard to operations because she did not possess a clearance at the time, the individuals I named are professionals and would not disclose classified information to an uncleared person. In any case, the AIR investigators cannot be excused from the attempt to suppress intellectual findings by, or to limit the research of, a noted academic that may be germane to the stated goals of the investigation.

The NRC study was discredited in print, and I had discussed that issue in detail with the AIR's blue ribbon panel.

Biased Investigators on the AIR's "Blue-Ribbon" Panel

Since our research program had been reviewed by various science advisory boards, including DIA's, it seemed prudent and natural that the CIA should ask their own Board or one of many that reside in the Washington area to conduct the program's technical evaluation. I even provided names and phone numbers of individuals whom I know on various boards to expedite the contact.

Instead, Utts and Hyman were chosen to act as the expert reviewers. At first glance, this seems like a reasonable approach given that no learning curve would be required. I told the POC that I thought this was not a good plan and that I could easily predict their conclusions based on their previous writing. (See Hyman, 1985, and Utts, 1991, as samples.) I reiterated that an in-place science advisory board would better serve that evaluation.

What better way to conclude whatever you wish than to build into the evaluation protocol a priori stated scholarly views that are known to span the opinion space. This guarantees that the concluding remarks by the CIA will, by definition, be consistent with some evaluator on the team. That is exactly what happened. In the CIA's presentation to Congress, eight separate bulleted points are allotted to Hyman's conclusion while only four are allotted to Utts's, and none are given to Utts's important rebuttal to Hyman.¹¹

¹¹This information was provided to me in 1995 by a U. S. Senate Appropriations Committee staff member.

Good Advice Ignored

Since most of the work under review occurred under my watch as the contractor program director, I could obviously not be involved in the analysis directly, but as part of my contract responsibility I was asked to advise the review process. In a four-page memorandum of June 6, 1995, to the POC, I indicated in words and figures how a review might proceed. The major point was that acceptance criteria for operations and research should be set prior to the review so that they could be used to judge the validity of the program in an unbiased way.

(Arguably, one could say that I had a vested interest in the outcome and that my views should be ignored. However, I only provided suggestions from a top-down perspective and did not suggest any details that could be considered self-serving. It was beneficial to the program and to me personally to have the most honest and rigorous review possible, and I was completely confident that such a review could only be positive.)

The criteria for the research could easily be adopted from the established and accepted scientific rules for evidence. In my June 6, 1995, memorandum, I stated:

The existence of anomalous mental phenomena cannot be statistically determined from the results of a single laboratory. The requirements for replication of a statistical phenomenon and the methods for the analysis of cross-laboratory results are well developed.

Not only was this advice ignored, it was ignored by fiat. The reviewers were instructed to only look at research results from SRI and SAIC. Fortunately for scientific credibility, Utts ignored this statistically invalid directive. Such action by the CIA with regard to their review can only add to the evidence that either they were only interested in a negative outcome or they were statistically naive.

Determining the efficacy of operations was much more difficult. Would one successful operation be sufficient justification to continue the program, or would all the operations have to be useful? What constitutes a successful operation? A 1% hit rate might be considered miraculous by one customer, but a 50% hit rate might be useless to another. I made no attempt to suggest what that judgment criterion should be; I only urged that it be set in advance. It wasn't.

It was not done as a matter of official policy or even informally as a guideline. As it turned out, the POC later informed me that only a single case would be sufficient as far as he was concerned, but he was careful to say that the decision was being made at "a much higher pay grade" than his. I learned later that they were going to examine only the last set of AC

operations from the 24-year program. I and they knew that these cases were not representative of the program at large. This point will be expanded below.

Early in the review, I was request to provide a list of my 10 best examples of research that supported the existence of AC. In a memorandum to the POC of June 14, 1995, I complained about that request:

Since the complete document set will be available to AIR, I recommend the following approach:

- For the period at SRI from 1973 to 1989 (this also covers the pre NRC report date) use the [in-house] meta-analysis as a guideline for the assessment with spot checks to the primary documents to validate the SRI evaluation.
- Use *all* the work conducted under the SAIC program from 1991 through 1994 as the simplified test set of documents. I think that includes 4 final reports and perhaps 10 major projects within that set.
- Conduct the final evaluation from both sources of data. (One thing that could be done is to use the results of the meta-analysis of the SRI data to predict what might happen during the SAIC research. The meta-analysis could be predictive only if there were a genuine phenomenon. In my view, this would add to the overall analysis.)

This approach avoids the file drawer problem [i.e., not publishing studies that fail to meet statistical significance] altogether and includes most of the documents I would count as my 10 anyway. I can only think of a few other studies that I might want to include and all of them have been accepted for publication in peer-reviewed journals.

In a memorandum to the POC of June 19, 1995, I responded in part again to the same request:

Although the request seems straight forward at the outset, to establish the existence of Remote Viewing on the basis of a subset of the total data set does not conform to the accepted practice for meta-analysis as set forth in Rosenthal (1991) and Hedges and Olkin (1985).

I went on to comply with the request in such a way that the complete record would be examined to avoid any accusation of a so-called "file-drawer" problem by including in my list a detailed in-house meta-analysis covering the period from 1973 to 1989 (May, Utts, Trask, Luke, Frivold, & Humphrey, 1989). This analysis was conducted as part of a contractual requirement to a still-classified sponsor.

The AIR ignored the CIA directive by including the NRC's review of parapsychology as a support for their conclusions about research. Knowing

full well that the NRC investigators did not have access to any SRI reports from 1985 onward,¹² they featured it prominently in their final report.

Little Contact with the Program's Principal Investigator

I would like to emphasize my role, or lack of it, in the CIA/AIR evaluation of the Star Gate program. As I said before, it was inappropriate for me to be involved in the actual assessment; however, it is especially important to learn from me the critical details that never made it into official reports.¹³ To illustrate my point: Of all the blue-ribbon panelists, Utts was the most familiar with the project; she had served as a visiting scientist for a year during the SRI era. Even with her intimate knowledge, she called me at least 12 times to seek clarification on specific points in the documents she was reading. Hyman never called me, and the AIR team not only did not call but refused to return my multi-faceted communication attempts. As a result of AIR negligence, their report contains numerous errors of fact and errors of assumption.

I was the director of the government-sponsored investigation of anomalous mental phenomena for 10 years of its 24-year history. I presided over 70% of the total contractor budget and 85% of the program's data collection, and I had intimate knowledge of and responsibility for the project. For the AIR not to use this resource is scientifically indefensible.

As the review process was coming to an end, I formally sought (in an August 14, 1995, memorandum to the POC) the opportunity to provide a written commentary to the AIR report to be included with the blue-ribbon panel's reports. Given that Utts and Hyman were given space to comment on each other's work,¹⁴ and since most of the science that was being reviewed was work conducted under my direction, it seemed only natural to include my comments. That request and a similar one to AIR were ignored.

¹²This was evident in my only meeting with the AIR blue-ribbon panel, in July 1995, when this lack of access was discussed in detail.

¹³There is no government or private program that can be accurately assessed exclusively from its written record.

¹⁴It is curious to note that the AIR report does not contain the original effort from Hyman but only his response to Utts.

POLITICAL REASON WHY THE CIA MAY NOT HAVE
WANTED THE PROGRAM

Under the reluctant auspices of the DIA, the program transitioned from SRI to Science Applications International Corporation in 1991. We recognized shortly thereafter that the DIA did not welcome the responsibility as the contracting agency. The reason that the DIA management was not anxious to have the program was complex and not associated with the technical aspects. Some of the DIA management had previous negative experiences with senior military officers who had become "believers," had oversold the program's capability, and were known as "loose cannons" in the community.

This reluctance manifested in two important ways. First of all, the initial financial support for the program in 1991 came directly as part of the supplemental Defense Appropriations bill and was considered by Congress as "seed" money. DIA was expected to request follow-on support as part of the overall DIA annual budget requests. Those requests never happened; all program support through 1995 came from the Appropriations bills. One consequence was that a member of the staff of the Senate Appropriations Committee became increasingly disappointed with the DIA and began to micromanage the program with disastrous results. A second consequence was that an attempt was made in 1993 to transfer the program to the CIA. No willing recipient could be found there to accept the program. Even then the CIA did not want program responsibility.

Secondly, the negative attitude from senior DIA management filtered down the chain of command. For example, the final project officer who had direct responsibility for the program before it closed had little knowledge of the program, no knowledge of its substantial history, and no technical background to manage such a project; and he ignored the research results and created a crushing atmosphere with his management style. The morale was so bad that viewers and officials within the government's remote-viewing unit repeatedly asked me to intervene. This placed me in a very difficult position as a contractor. I informed middle management at the DIA of the problems, with no result.

In short, the program was in shambles. The operations that were conducted during the last few years of the project were, for the most part, destined to fail and did so. It was this program, including personnel, that was to be transferred to the CIA by July 1, 1995. In my professional opinion, which I shared with the POC, the program as it was configured would not produce successful AC intelligence operations.

So the CIA had strong and valid reasons not to want the program. The CIA was soundly criticized in the press for mishandling the Ames

case and for other excesses, and so they did not need another controversy. In my opinion, the last thing they would want would be to inherit a congressionally micromanaged program in severe internal distress no matter what its content or potential might be. Yet by law they had to comply with the Congressional Directed Action and conduct the review. No wonder that it was probably done in such a way as to ensure a negative outcome with regard to operations.

CONCLUSIONS

It is impossible for me to prove whether or not the CIA determined the outcome before the investigation began. What is obvious, however, is that the evaluation of the research and particularly the operations were restricted so as to preclude positive findings. The CIA either did not contact or ignored people who possessed critical knowledge of the program, including some end-users of the intelligence data. Investigators were chosen who either had previously published conclusions or possessed a serious potential for a conflict of interest. With the exception of the significantly flawed NRC review, all the previous evaluations by the DOD of the research and intelligence applications were ignored. I am forced to conclude that either the AIR investigators were not competent to conduct a proper review of such a complex program—a view to which I do not subscribe—or they knew exactly what they were doing: They wanted to demonstrate that AC lacked usefulness for intelligence operations. They did so by construction rather than by careful analysis.

Let us grant for the moment that my supposition is true—that the CIA wanted to kill the program. Why was such a detailed rebuttal necessary? After all, an agency should be able to express their wishes with regard to the acceptance of any program that the Congress might assign. In fact, I see it as part of an agency's responsibility to inform Congress of what might or might not be possible. Rejecting the Star Gate program on the basis of an incomplete and incorrect analysis not only creates a false legacy, it does not easily allow for other organizations in the public or private sector to assume responsibility for a new version of the program. Aside from setting the record straight, I felt obligated to show that, as the result of their flawed methodology, the CIA/AIR greatly underestimated the statistical robustness of the research results and significantly undervalued the potential for AC in intelligence operations.

REFERENCES

- DRUCKMAN, D., & SWETS, J. A. (Eds.) (1988). *Enhancing human performance*. Washington, DC: National Academy Press.
- DUNNE, B. J., & BISAHA, J. P. (1979). Precognitive remote viewing in the Chicago area: A replication of the Stanford experiment. *Journal of Parapsychology*, **43**, 17-30.
- HEDGES, L. V., & OLKIN, I. (1985). *Statistical methods for meta-analysis*. New York: Academic Press.
- HYMAN, R. (1985). A critical historical overview of parapsychology. In P. Kurtz (Ed.), *A skeptic's handbook of parapsychology* (pp. 3-96). Buffalo, NY: Prometheus.
- JAROFF, L. (1974, March 4). Boom times on the psychic frontier. *Time*, 56-72.
- MAY, E. C., UTTS, J. M., TRASK, V. V., LUKE, W. L. W., FRIVOLD, T. J., & HUMPHREY, B. S. (1989). *Review of the psychoenergetic research conducted at SRI International (1973-1988)*. Final Report—Task 6.0.1, Project 1291. Menlo Park, CA: SRI International.
- MUMFORD, M. D., ROSE, A. M., & GOSLIN, D. A. (1995, September 29). *An evaluation of remote viewing: Research and applications*. Washington, DC: American Institutes for Research Report.
- OFFICE OF TECHNOLOGY ASSESSMENT. (1989). Report of a workshop on experimental parapsychology. *Journal of the American Society for Psychical Research*, **83**, 317-339.
- PALMER, J. A., HONORTON, C., & UTTS, J. (1989). Reply to the National Research Council study on parapsychology. *Journal of the American Society for Psychical Research*, **83**, 31-49.
- ROSENTHAL, R. (1990). Letter to the editor. *Psychological Science*, **1**, 329.
- ROSENTHAL, R. (1991). *Meta-analysis procedures for social research*. London: Sage Publications.
- SCHLITZ, M. J., & GRUBER, E. (1980). Transcontinental remote viewing. *Journal of Parapsychology*, **44**, 305-317.
- SCHLITZ, M. J., & HAIGHT, J. (1984). Remote viewing revisited: An intrasubject replication. *Journal of Parapsychology*, **48**, 39-49.
- TARG, R., & PUTHOFF, H. E. (1974a). Information transmission under conditions of sensory shielding. *Nature*, **252**, 602-607.
- TARG, R., & PUTHOFF, H. E. (1974b). Geller: Experimenters reply [letter]. *New Scientist*, Nov. 7.
- TARG, R., & PUTHOFF, H. E. (1977). *Mind-reach*. New York: Delacorte Press.
- TARG, R., KATRA, J., BROWN, D., & WIEGAND, W. (1995). Viewing the future: A pilot study with an error-detecting protocol. *Journal of Scientific Exploration*, **9**, 367-380.
- UTTS, J. (1991). Replication and meta-analysis in parapsychology. *Statistical Science*, **6**, 363-403.

Cognitive Sciences Laboratory
330 Cowper Street, Suite 200
Palo Alto, CA 94301