

The American Institutes for Research Review of the Department of Defense's STAR GATE Program: A Commentary

EDWIN C. MAY

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

Abstract — As a result of a Congressionally Directed Activity, the Central Intelligence Agency 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 was contracted to conduct the review of both research and operations. Their 29 September 1995 final report was released to the public 28 November 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 DOD extensive program reviews, and by using the discredited National Research Council's investigation of parapsychology as the starting point for their review. While 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, this 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 significantly underestimated the robustness of the basic phenomenon.

Executive Summary

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. 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. In addition, the CIA was to include an assessment of the intelli-

¹Hereafter we 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.

gence-gathering utility of anomalous cognition (AC), and the program history was to be declassified (CIA Public Affairs Office, 1995). Initiated in June 1995, the evaluation was to be completed by 30 September 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 Professor Ray Hyman from the University of Oregon and statistician Professor Jessica Utts from the University of California at Davis. AIR contributed Michael Mumford, Ph.D. and Andrew Rose, Ph.D. to the panel to provide unbiased assessment on methodological issues. The President of AIR, David Goslin, Ph.D., served as coordinator of the research effort.

I was asked by 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 Science Applications International Corporation.²

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

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

- Limited the data sets in the analysis. As a way of officially ignoring anomalous cognition'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.
- 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 project's reports 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:

- Apply consistent criteria for acceptance or rejection of anomalous cognition. The investigators were troubled by possible non-AC alternative

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

explanations for the statistically significant laboratory results, yet 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.

- 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 four months to complete the evaluation, it is surprising that they chose not to use this resource.
- 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.
- 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.
- To recognize a potential significant conflict of interest for Dr. 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 pre-determined perspective. In addition, I will document that as a result of their minimum effort, they have come to the wrong conclusion with regard to the use of anomalous cognition 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 only include government-sponsored research. If the evaluation was 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 if there were merit to continue.

Upon closer inspection, however, even in this case, the approach 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). Science demands that the evidence rest in replications; yet, the reviewers were requested not to look outside the STAR GATE project.

In their briefing to Congress, the CIA list three points as attributed to the AIR investigation (May, 1995g):

- "the data do not establish that a paranormal phenomenon is involved, nature of source not identified"
- "the data have not been replicated independently"
- "the boundary constraints critical to obtaining statistically significant experimental results are not practical in the 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 bullets above are true by construction — not by analysis.

As an aside, 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, *et al.*, 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 were designed into the investigation. However, their final bullet above is equally absurd on its face value, because it is true by the nature of intelligence, not because of a valid criticism of the program's operational AC. The only valid measure of intelligence utility for anomalous cognition 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 they can point to that helped solve an intelligence problem? The CIA and AIR say no, but as I will show below, that conclusion also was 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 anomalous cognition 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 three years from the end of the 24-year program. I argued strenuously with him because 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 based on my knowledge of the morale of the government's viewers, the substandard 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. I informed him that 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. In an urgent call to the individual at CIA who had been assigned to manage the review, hereafter called the Point of Contact or POC, I insisted that some of the officials I named previously 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, both on active duty and retired, had already been on written record as attesting to the value of AC-derived intelligence data in solving individual 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 were not contacted and those that were told the CIA representative the case specifics and value of their individual circumstances. Some of the positive findings occurred before the final year but within the last three years of the project; perhaps that is why the "official" investigation only went back a single year in spite of Mumford's original plan to look at the last three years.

Finally, even a cursory investigation of the written record of intelligence operations would have revealed substantial evidence of the operational utility of anomalous cognition. 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, was granted a Legion of Merit award in 1984 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 the scientific reputations and on the basis of their honest skepticism. "Believers" were not allowed on the committee. The SOC's responsibilities were four-fold:

- Evaluate our written protocols prior to conducting any experiments. The protocol that was actually used for each investigation was the convergence of a round-robin exchange with the SOC.
- Exercise unannounced drop-in privileges to observe experiments in progress. Approximately one half of the SOC availed themselves of this opportunity.
- Review the then-classified final research reports as if they were technical journal submissions in their individual disciplines. The disciplines included physics, philosophy, psychology, electrical engineering, statistics, and astronomy. Their reviews were in writing and appended, unedited, to each final report.
- 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. The list included one Nobel laureate as did SAIC's Scientific Oversight Committee. 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 CIA and that was done after strenuous insistence on my part. It was that single individual who provided the names of satisfied end-users I discussed above.

The SOC's comments were available to the AIR reviewers in written form, but given that many of the committee members live on the east coast and even a few live in Washington it is astonishing to me that they were never contacted personally. It is simply not possible to obtain an honest evaluation of our project without talking to these overseers. It's ironic that the government spent

considerable resources assuring the highest possible technical and administrative oversight specifically to provide on-going and independent assessment of the program; yet, CIA/AIR chose to ignore it.

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 former 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 to the scientific, operational, managerial, and political aspects of the STAR GATE project and in particular, its potential. 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.* (p. 2-1, 1995) correctly required for 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 anomalous cognitive one, although he admits he can'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 effect the result, he is correct in proposing the hypothesis that it might somehow affect 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. As an aside, this same consistency also rejects the other alternatives Hyman proposes, as well.

Yet, AIR fails 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 on Evaluating the Utility of Remote Viewing in Intelligence Operations they list a number of "boundary conditions" that might affect anomalous cognition in operations. These include a number of physical and methodological issues such as feedback and whether a sender or distance of the target might be factors.

What is surprising to me is that 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, the expectations of the participants, and 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 two 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 (May, 1995i) The morale was so low that I told the panel that I couldn't have balanced my check book correctly in that environment; doing excellent remote viewing 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 (May, 1995j).

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 by the Government

Even before I was officially under contract with CIA, I gave the POC either copies of, or pointers to, a number of classified program reviews that had been conducted in the past.³

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 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 only wanted 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 observation 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.

³There have been a number of unclassified reviews and published meta-analyses of anomalous mental phenomena, but they do not deal with operations.

At the outset, I was excited about the potential of a fair and honest evaluation of the complete program, and was delighted that the new review could essentially encompass the work conducted since 1987. The POC was a Ph.D. scientist and seemed dedicated to the best job possible. He informed me, however, that they intended to ignore the previous reviews and start fresh. I was shocked. Given that the review had to be in Congress in four 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.

It is only from the perspective of a pre-determined outcome that such a policy could be understood.

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 Mr. 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 is easy to become quite paranoid about who is trying to trick whom. The safest and the most scientifically sound course is not to allow anyone except the direct research team to witness formal experiments regardless of credentials (Targ and Puthoff, 1977 and May, 1996).

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 (Targ and Puthoff, 1974b and May, 1996). 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 American Academy of Sciences to investigate the potential of certain techniques that propose to enhance human performance (Druckman and Swets, 1988). Although it has never been the claim of research parapsychology, the National Research Council 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 Sci-

ences and Education for the National Research Council, served as overall project director and agreed to the request.

On 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 at 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.

By the time the NRC began its 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 clearances⁴ Thus, the NRC investigators and Ray Hyman had access to a small fraction 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 the this group, he was denied access to and probably even unaware of the SRI data of that time period.

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 1979.⁵

In the overall summary shown above, remote viewing was not even mentioned although an analysis of the early studies at SRI and later studies at Princeton are contained in the body of the NRC report. With regard to their conclusion on remote viewing: "...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."

The parapsychology section of the NRC study was a mockery of good science and serves as an excellent model for a pseudo-scientific investigation. The methodology for the NRC investigation and their conclusions were soundly criticized and shown to be without scientific merit (Palmer *et al.*, 1989). The four major points drawn by Palmer *et al.* are summarized:

⁴The research from 1985 through 1990 was finally declassified as a result of my petition to do so in 1990. Many of the sponsors and most of the intelligence operations, however, remain classified.

⁵It 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).

"The NRC claimed they could find no evidence for parapsychological phenomena during the last 130 years, yet they examined only 10% of the systematic scientific effort in parapsychology."

- "The two principal evaluators of parapsychological research, Ray Hyman and James Alcock, were publicly committed to a negative position on parapsychology at the time the NRC Research Committee was formed. [Note added by May: In addition, the phrase "...the total accumulation of 130 year's worth of psychical investigations has not produced any consistent evidence for paranormality..." can be found in both Hyman (1986) and the NRC conclusion (1988), and thus demonstrates his stated bias before the NRC investigation was complete.]"
- "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 and Swets, 1988, p. 200).
- "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."

This last point is particularly heinous and reveals the political nature of what should have been a carefully conducted scholarly investigation that usually characterizes the National Research Council. Violating one of the basic tenets of science to report all findings, the NRC Committee asked Professor 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.⁶

Rosenthal's and Harris' commissioned paper listed the Ganzfeld methodological quality to be superior to the typical quality of the other four areas they considered (Rosenthal, 1990), but this conclusion was not included in the NRC report.

In addition to the significant methodological flaws and the attempt to suppress positive findings, the NRC study was essentially contradicted in its major conclusion by a one-day workshop hosted by the Office of Technology

⁶Quoted with permission from a letter from Professor Rosenthal.

Assessment, the research arm of the US Congress (Office of Technology Assessment, 1989). The OTA did not completely exonerate the field of research 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 continuing 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 its technical review of the STAR GATE program, they were turned down by the National Research Council in part because of the time constraint and in part because of the substantial negative publicity that resulted from their previous report on parapsychology (May, 1995e). 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 had a long history of a negative bias with regards 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, and hired Ray Hyman as part of the evaluation team.

It is not a surprise to me that the NRC study is liberally quoted in the AIR report, because I believe it supports the predisposed views of CIA/AIR, albeit from a totally flawed investigation. Since Professor Jessica Utts was one of the co-authors of the formal response to the NRC study, I questioned her (May, 1995f):

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 to have added your first-hand criticism of the NRC report as part of your remarks.

Professor Utts' answer shocked me. So that I make no errors in interpretation, I print, with permission, her complete reply (19 December 1995):

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."

While a case can be made that Professor 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. Regardless, 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 AIR's blue ribbon panel. The fact that AIR liberally used the flawed

NRC investigation clearly demonstrates that a pre-determined negative outcome was in place.

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 who I know on various boards to expedite the contact; after all we were on a tight schedule.

Much to my dismay, 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. By phone, 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 (1986) and Utts (1991) as samples. I reiterated that an in-place Science Advisory Board would better serve that evaluation. After being told that they were moving ahead with Hyman and Utts as the evaluators, it became clearer to me that a set-up was in progress.

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 CIA will, by definition, be consistent with at least one 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' and none are given to Utts' important rebuttal to Hyman (May, 1995g).

Good Advice Ignored

Since most of the work under review occurred while I was the contractor program director, I could obviously not be involved in the analysis directly, but as part of my CIA contract responsibility, I was asked to advise the review process. In a 4-page document (May, 1995a), 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 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 would have been 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. Quoting from my memorandum (May, 1995a):

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

It was not done as a matter of official policy nor 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 only going to examine 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. The CIA continued to set up the operational review to fail.

Early in the review, I was request to provide a list of my 10-best examples of research that supported the existence of anomalous cognition. In a memorandum (May, 1995b), I complained about that request. In part, I quote:

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 draw 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 in-

clude and all of them have been accepted for publication in peer-reviewed journals.

I responded in part again to the same request (May, 1995c):

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 to 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 *et al.*, 1989). This analysis was conducted as part of contractual requirement to a still-classified sponsor.

AIR ignored the CIA directive not to consult previous assessments by including the National Research Council'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 (May, 1995d), they featured it prominently in their final report.

This piece of obvious subterfuge could only have happened in order to support a predetermined outcome that was known to be false.

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 for me to provide the context and critical details which often do not make it into official reports.⁷ To illustrate my point, of all the "blue-ribbon" panelists, Professor 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. Professor Hyman never called 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 assumptions.

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

As the review process was coming to an end, I formally sought the opportunity to provide a written commentary to the AIR report to be included with the

⁷There is no government or private program that can be accurately assessed exclusively from its written record.

blue-ribbon panel's reports (May, 1995h). Given that Utts and Hyman were given space to comment on each others work,⁸ 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 was ignored. It is my guess that since they probably set up the review to fail with regard to operations, and they knew how much I knew about the program history, they could not afford to learn officially even a fraction of what I have documented in this paper.

Political Reason Why CIA May Not Have Wanted the Program

Under the auspices of the DIA, the program transitioned from SRI to Science Applications International Corporation in 1991. We recognized shortly thereafter that DIA did not welcome the responsibility as the contracting agency. The reason 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 uncritical fanatics, 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. A member of the staff of the Senate became increasingly disappointed with the DIA and began to micro-manage the program with disastrous results, and an attempt was made in 1993 to transfer the program to 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, during the last two years of the program, DIA assigned a project officer who had no training or experience for the job. Most importantly, he ignored proper operational protocols and ignored the viewers' attempts at educating him. In addition, the project officer had little knowledge of the program's extensive history, nor did he possess the technical background to manage such a program. 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 DIA of the problems with no result.

In short, by 1995 the program was in shambles. The operations that were conducted during the last few years of the project, for the most part, were destined to and did fail. It was this program, including the personnel, that was to

⁸It is curious to note that the AIR report does not contain the original effort from Hyman but only his response to Utts.

be transferred to CIA by 1 July 1995. In my professional opinion, which I shared with CIA's POC, the program as it was configured would not produce successful AC intelligence operations.

So, CIA had strong and valid reasons not to want the program. The Agency was soundly criticized in the press for mishandling the Ames case and other excesses, so they did not need another controversy. In my opinion, the last thing they would want would be to inherit a Congressionally micro-managed 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 done in such a way to assure a negative outcome with regard to operations.

Conclusions

It is impossible for me to prove whether or not the CIA determined the outcome of the investigation before it began. What is obvious, however, is that the evaluation domain of the research and particularly the operations were restricted to preclude positive findings. The CIA 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 who possessed a serious potential for a conflict of interest. With the exception of the significantly flawed National Research Council's review, all of the DOD's previous evaluations of the research and intelligence application 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 a lack of intelligence utility for anomalous cognition. They did so by construction rather than by careful analysis.

Let us grant for the moment that my supposition is true, 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 the various agencies' 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 seriously flawed methodology, the CIA/AIR greatly underestimated the statistical robustness of the research results and significantly undervalued the potential for anomalous cognition in intelligence operations.

References

- Druckman, D. and Swets, J. A. Ed. (1988). *Enhancing human performance*. Washington, D.C.: National Academy Press, 167.
- Dunne, B. J. and Bisaha, J. P. (1979). Precognitive remote viewing in the Chicago area: A replication of the Stanford experiment. *Journal of Parapsychology*, 43, 1, 1.
- Hedges, L. V. and Olkin, I. (1985). *Statistical methods for meta-analysis*. New York, NY: Academic Press, Inc. Harcourt Brace Javanovich.
- Hyman, R. (1986). Parapsychological research: A tutorial review and critical appraisal. *Invited Paper. Proceedings of the IEEE*, 74, 6, 825.
- Jaroff, L. (1974). Boom times on the psychic frontier. *Time Magazine*, 4 March, 56.
- May, E. C., Utts, J. M., Trask, V. V., Luke, W. L. W., Frivold, T. J., and 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.
- May, E. C. (1995a). Memorandum to POC, 6 June 1995.
- May, E. C. (1995b). Memorandum to POC, 14 June 1995.
- May, E. C. (1995c). Memorandum to POC, 19 June 1995.
- May, E. C. (1995d). In May's only meeting with the AIR "blue-ribbon" panel during which the lack of NRC access was discussed in detail. July 1995.
- May, E. C. (1995e). Personal communication. The CIA point of contact.
- May, E. C. (1995f). E-mail letter to Professor Utts, 17 December.
- May, E. C. (1995g). Personal communication. US Senate Appropriations Committee Staff Member.
- May, E. C. (1995h). Memorandum to POC, 14 August 1995.
- May, E. C. (1995i). Personal communication. Foreign analyst for the unit, June, 1995.
- May, E. C. (1995j). Personal communication. A government remote viewer, August 1995.
- May, E. C. (1996). Personal communication. Hal Puthoff.
- Mumford, M. D., Rose, A. M., and Goslin, D. A. (1995). An evaluation of remote viewing: Research and applications. The American Institutes for Research Report, September 29.
- Office of Technology Assessment (1989). Report of a workshop on experimental parapsychology. *Journal of the American Society for Psychical Research*, 83, 4, 317.
- Palmer, J. A., Honorton, C. and Utts, J. (1989). Reply to the national research council study on parapsychology. *Journal of the American Society for Psychical Research*, 83, 1, 31.
- Public Affairs Office, Central Intelligence Agency, 6 September 1995.
- Puthoff, H. E. and Targ, R. (1976). A perceptual channel for information transfer over kilometer distances: Historical perspective and recent research. *Proceedings of the IEEE*, 64, 3, 329.
- Puthoff, H. E. (1996). CIA-Initiated remote viewing program at Stanford Research Institute. *Journal of Scientific Exploration*, 10, 1, 63.
- Rosenthal, R. (1990). Letter to the Editor. *Psychological Science*, 1, 5, 329.
- Rosenthal, R. (1991). *Meta-analysis procedures for social research*. London, UK: Sage Publications.
- Schlitz, M. J. and Gruber, E. (1980). Transcontinental remote viewing. *Journal of Parapsychology*, 44, 4, 305.
- Schlitz, M. J. and Haight, J. (1984). Remote viewing revisited: An intrasubject replication. *Journal of Parapsychology*, 48, 1, 39.
- Targ, R. and Puthoff, H. E. (1974a). Information transmission under conditions of sensory shielding. *Nature*, 252, 602.
- Targ, R. and Puthoff, H. E. (1974b). Geller: experimenters reply. *New Scientist*. Letters. 7 November.
- Targ, R. and H. E. Puthoff (1977). *Mind reach*. New York: Delacorte Press.
- Targ, R., Katra, J., Brown, D., and Wiegand, W. (1995). Viewing the future: A pilot study with an error-detecting protocol. *Journal of Scientific Exploration*, 9, 3, 367.
- Utts, J. (1991). Replication and meta-analysis in parapsychology. *Statistical Science*, 6, 4, 363.

Precognitive Remote Perception: Replication of Remote Viewing

R. D. NELSON, B. J. DUNNE, Y. H. DOBYNS, R. G. JAHN

*Princeton Engineering Anomalies Research, School of Engineering and Applied Science, C-131,
Princeton University, NJ 08544*

The following brief description of the Princeton Engineering Anomalies Research (PEAR) Remote Perception program has been prepared at the invitation of the Editor¹, in order to augment this special report section of the Journal with information about another substantial database of experiments relevant to those of SRI and SAIC. Given Utts' attention to the importance of replication (Section 3.4), and Hyman's challenge of interlaboratory consistency (Point #3 of his Introduction and Point #2 of his "Suggestions for Future Research"), we submit that the PEAR program has obtained the largest extant body of experimental data that meets their criteria for interlaboratory replication. In point of fact, both the PEAR remote perception program, and the prior studies of Dunne and Bisaha on which it was originally based, were undertaken as formal replications of the SRI experiments of Puthoff and Targ.

Although the PEAR program has accumulated several hundred experimental trials, its primary goal has been to develop a sophisticated analytical judging methodology to replace the human judging process, and thereby to facilitate more precise quantitative assessment of results and their correlation with various experimental parameters. In our basic procedure, the "free response" of the percipient is encoded using a list of 30 binary descriptor questions, allowing algorithmic comparison with the target, similarly encoded by the agent at the scene. For randomly assigned targets, further comparison can be made with an encoding by the person who prepared the target pool. The analysis proceeds by constructing a square matrix of scores calculated by comparing each perception against all targets in the given dataset. The properly matched trials (on the main diagonal of the matrix) can be assigned statistical merit by comparison with the distribution of off-diagonal, mismatched scores, which has sufficiently Gaussian characteristics to allow robust parametric statistical tests.

Beyond the primary experimental question of the degree of anomalous acquisition of information, several other issues have been explored, among them the correlation of analytical and human judge scores, the efficacy of different scoring algorithms and descriptor sets, *ex post facto* vs. participant encoded descriptions, agent chosen versus randomly assigned targets, single vs. multi-

¹Editor's Note: To be followed by a detailed, peer-reviewed article in the future.

ple percipients, variations among individual agent and percipient pairs, and the relationship of scores to the distance and time intervals separating the perception and the target.

The results in all phases of this experimental program are quite consistent with those of their SRI predecessors and with the more contemporary SAIC studies. Overall they show average effect sizes well within the range described by Utts (Sections 3.4, 4.2, and 4.3). For example, for the entire 336 trials comprising the formal PEAR database, the effect size (composite *Z*-score normalized by the square root of the number of trials) is 0.347 ± 0.055 . When these are separated into randomly assigned vs. volitionally chosen target subsets, the 125 randomly assigned targets show an effect size of 0.516 ± 0.089 , and the 211 volitional targets an effect size of 0.244 ± 0.069 . Assessment of individual performance indicates that the overall yield is an accumulation of small contributions from the majority of the participant pairs, rather than from a few outstanding efforts.

Among the more interesting findings is parametric evidence that the degree of anomalous information transfer is unaffected by spatial and temporal separations. Regression modeling indicates a significant mean shift, but no evidence for a decline of scoring with increasing distance, up to several thousand miles. Similarly, there is no evidence that scoring is related to positive or negative temporal separations of the perception effort and the target visit, up to as much as a few days. The precognitive subset of these data, consisting of about 75% of the independent trials, seems particularly important to the postulation of viable theoretical models, and has been emphasized throughout.

Thus, these databases, comprising one of the largest accumulations of relevant experiments performed under consistent and well controlled experimental protocols, have already provided robust evidence that the findings in the SRI/SAIC Remote Viewing experiments can be replicated in independent, but essentially similar designs. For more details, consult the following references:

- Dobyns, Y. H., B. J. Dunne, R. G. Jahn, R. D. Nelson (1992). Response to Hansen, Utts, and Markwick: Statistical and methodological problems of the PEAR remote viewing experiments. *J. Parapsychology*, 56, 2, 115.
- Dunne, B. J. and J. P. Bisaha (1979). Precognitive remote viewing in the Chicago area: a replication of the Stanford experiment. *J. Parapsychology*, 43, 1, 17.
- Dunne, B. J., R. G. Jahn, R. D. Nelson (1983). Precognitive remote perception. *PEAR Tech. Report 83003*.
- Dunne, B. J., Y. H. Dobyns, and S. M. Intner (1989). Precognitive remote perception III: Complete binary database with analytical refinements. *PEAR Tech. Report 89002*.
- Jahn, R. G. (1982). The persistent paradox of psychic phenomena: An engineering perspective. *Proceedings IEEE*, 70, 2, 136.
- Jahn, R. G. and B. J. Dunne (1986). On the quantum mechanics of consciousness, with application to anomalous phenomena. *Foundations of Physics*, 16, 8, 721.
- Jahn, R. G. and B. J. Dunne (1987). *Margins of reality: The role of consciousness in the physical world*. New York, Harcourt Brace.
- Jahn, R. G., B. J. Dunne, R. D. Nelson (1987). Engineering anomalies research. *Journal of Scientific Exploration*, 1, 1, 21.
- Jahn, R. G., B. J. Dunne, E. G. Jahn (1980). Analytical judging procedure for remote perception experiments. *J. Parapsychology*, 44, 3, 207.