Approved For Release 2000/08/08 : CIA-RDP96-00789R002100030001-5

SECRET

DT-S-1010-S

DEFENSE
INTELLIGENCE
AGENCY

EVALUATION METHODS (U)

INITIAL REPORT

13 December, 1990

NOFORN

SECRET

Approved For Release 2000/08/08 : CIA-RDP96-00789R002100030001-5

Approved For Release 2000/08/08 : CIA-RDP96-00789R002100030001-5

[redacted]

EVALUATION METHODS

Date of Publication
13 December 1990

This is a Department of Defense Intelligence Document prepared by the Technology Assessment and Support Office, Directorate for Scientific and Technical Intelligence Defense Intelligence Agency.

PREPARED BY
Technology Assessment and Support Office (DT-S)

NOT RELEASABLE TO FOREIGN NATIONALS

Classified by: DIA/DT
Declassify: OADR

[redacted]

Approved For Release 2000/08/08 : CIA-RDP96-00789R002100030001-5

Approved For Release 2000/08/08 : CIA-RDP96-00789R002100030001-5

[redacted]

TABLE OF CONTENTS

[redacted]

Approved For Release 2000/08/08 : CIA-RDP96-00789R002100030001-5

Approved For Release 2000/08/08 : CIA-RDP96-00789R002100030001-5

EVALUATION METHODS

I. (U) PURPOSE:

(S) The purpose of this initial report is to provide a summary of methods that have been used for evaluating psychoenergetics (ie., remote viewing) data in both a research and applications-oriented environment.

II. (U) BACKGROUND:

(U) When modern research into psychoenergetics (ie. Extrasensory Perception) began in the early 1930's at Duke University, NC., the experiments were designed to accommodate easy to use statistical methods. Consequently, a small number of forced choice targets (e.g., a set of five cards with different symbols) were developed as targets. Results from experiments using such targets could be readily compared with those expected from chance guessing. If results exceeded a preset value (usually one out of twenty) a case for phenomenon existence could be made, especially if the experimental trials were large in number (ie., several thousand).

(U) Although these early statistical methods were convenient, they could not be applied to evaluate results from the remote viewing experiments that began in the early 1970's. Targets in most remote viewing experiments are not limited to a small set of possibilities; most early remote viewing experiments

Approved For Release 2000/08/08 : CIA-RDP96-00789R002100030001-5

Approved For Release 2000/08/08 : CIA-RDP96-00789R002100030001-5

used natural sites or National Geographic pictures that could be almost anything. In addition, data output was radically different from that required for early ESP research. Instead of "guesses" as to what card was the target (usually, a hunch or intuition prompted the participant), a remote viewing subject actually developed pictorial and written material. Data from remote viewing sessions required a "free response" style that by its very nature did not fit with any clear-cut statistical approach. Consequently data evaluation, as well as phenomenon existence assessments, became much more complex.

(U) Several statistical-based methods were subsequently developed to help in remote viewing data assessment. Over-all results, even if statistically significant, could not make a strong case for phenomenon existence due to the small number of trials involved in a typical remote viewing series. It was inherently more time consuming to perform a single remote viewing experiment than a card guessing series of hundreds, if not thousands, of trials for any single participant.

(S) Consequently, free response evaluations from the research environment were aimed at assessing data uniqueness on a trial-to trial basis. These required establishment of large (at least 100 or more) homogeneous targets in a fixed target pool that needed complex statistics for assessing results. Improved evaluation methods followed based on artificial intelligence approaches. Unfortunately, most if not all of these statistical

Approved For Release 2000/08/08 : CIA-RDP96-00789R002100030001-5

Approved For Release 2000/08/08 : CIA-RDP96-00789R002100030001-5

approaches were difficult to apply in an operational environment where targets were not homogenous and where importance of various target elements varied from one trial (or project) to another. Worse, not all relevant information required for data assessment would be known, and some "ground truth" data may even have been wrong.

(S) Nonetheless an evaluation procedure, even if subjective in nature, had to be developed initially to at least provide some basis for estimating or evaluating the results from operational projects. Later, improved methods that examined both data accuracy and data reliability were developed based on methods used for applications research projects. These improved methods have reduced, but not eliminated, the subjective nature of operational project evaluations. The accumulation of a large track record for given individuals over time, and performing meta-analysis of this accumulated data base, would be needed to further improve over-all assessment of such remote viewing data.

(S) Even though workable methods for assessing data accuracy and reliability have been developed, there is yet another consideration for operational projects: How useful was the data? This requires another set of evaluation (ie., utility assessment) that is customer dependent. Results from such assessments, unfortunately, ranged widely due to differences in evaluation criteria that were used. Utility assessment criteria need to be defined in advance of any operational project in order to minimize

Approved For Release 2000/08/08 : CIA-RDP96-00789R002100030001-5

Approved For Release 2000/08/08 : CIA-RDP96-00789R002100030001-5

subjective aspects. This has not always been accomplished in the past and needs to be part of any future application project procedure, whenever possible.

(C) However, specific evaluation methods are only meaningful when the over-all remote viewing activity is based on sound methodological procedures. Procedural aspects can easily be developed to ensure that the activity is in fact consistent with sound scientific methodology. Appropriate procedures are discussed in companion reports, and are only briefly addressed in this report.

III. (U) SCOPE:

(C) The following sections provide brief summaries of the various approaches that have been used by this unit for evaluating results from operational or applications-oriented projects. A follow up report is planned that will review in more detail the evaluation methodologies used for research and for applications-oriented projects. Other relevant issues are also discussed.

IV. (U) OPERATIONAL PROJECT EVALUATION METHODOLOGIES:

(S) There are two main issues in evaluating remote viewing data; (1) What is the definition of the target; (2) What is the definition of the remote viewing response. Various methods examined are simply different ways of comparing and evaluating the target and the response.

Approved For Release 2000/08/08 : CIA-RDP96-00789R002100030001-5

A. (U) SCALE APPROACH:

(S/NF) A subjective based value scale of 0 through 5 was used in the past; a value of 0 indicated no correlation to ground truth; a value of 5 indicated a perfect match. Recently, scale values of 0 through 3 (with +, -, variations) have been used. By whatever range of scale values used, the viewers' raw or summarized data is compared to known information about a target. The best possible judgement is made concerning approximate degree of correlation to "ground truth". An example of a specific scale evaluation approach is shown on figure 1. This 0 through 3 evaluation scale illustrates numerical ratings, percentages, and descriptions for degree of correlation with regard to essential elements of information (EEI) desired for each project.

(S/NF) Figure 2 lists the major target catagories that are usually of interest in any remote viewing project. Not all of these are of concern for any given task. Complex targets such as S&T facilities for example, are generally more difficult to evaluate than straight forward projects which have only 1 or 2 elements to consider. Where possible, major target catagories of interest (e.g., facility function) would be specified as part of the desired information in advance of the session. However, all the raw data is examined no matter what its relative importance or category. This provides a gauge of individual strengths and weaknesses useful for future target/person matching.

EVALUATION SCALES

NUMERICAL RATING	APPROX. DEGREE OF DATA CORRELATION PERCENT	DESCRIPTION
0	0-10	LITTLE OR NO CORRELATION
1	10-30	MIXTURE OF RELEVANT AND INCORRECT DATA WITH MAJORITY INCORRECT
1+	30-50	MIXTURE OF RELEVANT AND INCORRECT DATA WITH MAJORITY INCORRECT
2	50-70	MIXTURE OF RELEVANT AND INCORRECT DATA WITH MAJORITY CORRECT AND UNAMBIGUOUS
2+	70-90	MIXTURE OF RELEVANT AND INCORRECT DATA WITH MAJORITY CORRECT AND UNAMBIGUOUS
3	90-100	VERY HIGH TARGET CORRELATION WITH ESSENTIALLY NO AMBIGUOUS DATA

FIGURE 1

Approved For Release 2000/08/08 : CIA-RDP96-00789R002100030001-5

DATA CATEGORIES

GEOGRAPHIC DESCRIPTIONS -- **e.g., TERRAIN FEATURES; WATER, MOUNTAINS**

OBJECTS (LARGE SCALE) -- **e.g., AIRFIELDS, FACILITY LAYOUT, DOCKS, SILOS**

OBJECTS (SMALL SCALE) -- **e.g., LASERS, TANKS, ANTENNAS**

FUNCTIONS (LARGE SCALE) -- **e.g., R/D, WEAPON TESTING, PRODUCTION**

FUNCTIONS (SMALL SCALE) -- **e.g., LASER TESTING, CBW STORAGE, MISSILE FIRING**

PERSONALITY DATA -- **e.g., STATE-OF-HEALTH, PHYSICAL FEATURES, BACKGROUND**

PREDICTIVE -- **e.g., PLANS, INTENTIONS, FUTURE LOCATIONS, FUTURE ACTIONS/ACTIVITY**

FIGURE 2

Approved For Release 2000/08/08 : CIA-RDP96-00789R002100030001-5

Approved For Release 2000/08/08 : CIA-RDP96-00789R002100030001-5

B. (U) FUZZY SET APPROACH:

(C) Fuzzy Set Theory is an objective mathematical framework for both verbal and visual analysis which is utilized for evaluating imprecise data. Imprecision results from sketches that illustrate general shapes or approximate spatial relationships. Verbal data generally includes more content (ie., analysis) than visual information. The fuzzy set theory allows for numerical values to be assigned to target elements that represent their degree of importance. These numerical values would be assigned by consumers (or other expert personnel) in advance of any project. Numerical estimates are also made of the raw data after the session that represents its degree of correlation to and importance with the intended target. Thus, the remote viewing data can then be quantified by appropriate calculations to determine data accuracy and reliability. Accuracy is defined as the percentage of target material that is described correctly by a viewer; reliability is defined to be the percentage of the over-all response that correlates to the target.

(S/NF) Figure 3 shows an example of some of the data provided in a recent remote viewing experiment conducted by Stanford Research Institute to illustrate this procedure. In this example, the target used for the experiment was a microwave generator, support equipment, and testing equipment. A viewer described over seventy functions, objects and relationships. Over-all accuracy,

Approved For Release 2000/08/08 : CIA-RDP96-00789R002100030001-5

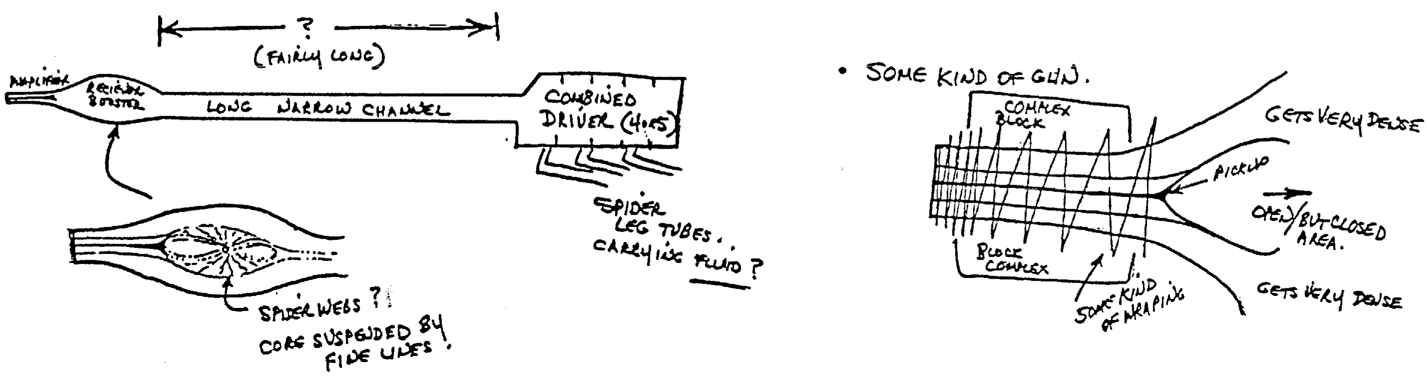
using the fuzzy set approach, was calculated to be about 80%. However, over-all reliability was only 65%; this indicated that 35% of the raw data had no correlation to the intended target. The product of accuracy and reliability yields a "figure of merit" what is also useful for over-all data assessment and for examining viewer performance over time. Additional details with regard to this experiment, are in reference no. 11 of the bibliography. In practice, a simplified version of this approach can be used to minimize analysis time.

C. (U) CONCEPT ANALYSIS:

(S/NF) Concept analysis is based on analyzing data according to the over-all concept rather than on smaller bits of information usually found in a remote viewer's response. For example, in figure 3, one of the responses to a target was "a fairly long narrow channel". In concept analysis, the concept of tube, or possibly gun, would be emphasized rather than breaking apart original words such as long, narrow, or channel. Although this can be a useful approach, some meaningful data may be over looked. This method had not been widely used, although it was useful for initiating the fuzzy set approach.

D. (U) CONTROL GROUPS:

(U) In Research & Development activities, control groups are often necessary to establish a data baseline to which the results of other experiments can be compared. Generally, a control group is a randomly selected group of people run according



SECRET

SECRET

SG1B

[redacted]

SECRET/NOFORN

REMOTE VIEWING DATA/TARGET EXAMPLE

FIGURE 3

Approved For Release 2000/08/08 : CIA-RDP96-00789R002100030001-5

to one set of protocols, while another group is run according to a different set of protocols. Both groups seek to achieve the same goal. After a designated testing period, the results are compared. The results indicate the effectiveness of each protocol independently of the other. Control groups can also indicate the extent to which remote viewing data is different from data generated by knowledgeable experts given the same background information. This can permit estimates to be made regarding validity of the remote viewing data based on standard statistical procedures.

E. (U) <u>IN-GROUP CONTROLS:</u>

(U) A simple method often used in the research community involves the use of a comparative approach. In this method, the raw data from a session is compared to one of several possible targets, one of which is the correct one. Judges blind to the actual target attempt to make the best match. If they succeed, a case can be made for remote viewing success. Statistics are straight forward. However, due to low numbers of targets generally used for this comparison (usually 4 to 6), statistical strength is quite low. This method is useful, however, and can provide insight into the remote viewing process. It does, however, minimize the significance of highly unique data elements and is not a good indication of data usefulness.

V. (U) <u>UTILITY EVALUATION CONSIDERATIONS:</u>

(S/NF) Utility refers to how useful remote viewing data

Approved For Release 2000/08/08 : CIA-RDP96-00789R002100030001-5

Approved For Release 2000/08/08 : CIA-RDP96-00789R002100030001-5

proved to be to the consumer. Usually some form of scale value or statements ranging from "very useful" to "of little use" have been provided as feedback. These utility evaluations, when performed, have usually suffered from lack of consistent evaluation criteria. In many cases, the criteria were not ever agreed upon in advance of project initiation. This is an important issue that will be addressed in future applications-oriented research.

VI. (U) <u>EVALUATION/DATA PROBLEMS:</u>

(S/NF) Sometimes it is difficult to complete evaluations since ground truth may not be totally known, or possibly the raw data may contain predictive information of an unspecified future time period. In such cases, only partial evaluations are possible, and final assessments may require months or years. Such potential delays in data evaluation pose serious problems for the reviewer (ie., feedback not possible), as well as for the consumer who may require timely information.

(S/NF) Another problem is who should do the evaluating? If only customers evaluate the operational data, they may not be capable of observing trends or patterns that could be useful. If evaluations are solely determined by individuals not involved in operations, they may emphasize aspects that are not operationally important. A combination of both views must be considered when possible and implemented in the evaluation process. It is also necessary to minimize or eliminate the role of the data procedures in final data evaluation since this would present a potential for

Approved For Release 2000/08/08 : CIA-RDP96-00789R002100030001-5

assessment biasing.

(S/NF) Other problems can include viewers being exposed to inferences and deductions by open press sources or television for certain current event projects. A monitor or observer present during a remote viewing session could also pose a problem since he or she could unknowingly bias the viewer. Biasing the source can be due to subliminal cueing. Consequently, thorough records must be kept regarding possible target related knowledge of those present in the remote viewing session.

VII. (U) <u>PREREQUISITES FOR EVALUATION:</u>

A. (U) <u>RESEARCH REQUIREMENTS:</u>

(S) Research requirements are more stringent than operational requirements since proof of principle or the search for difficult-to-detect variables are involved. Consequently, there is a strong need for well-defined targets and tightly controlled protocol so that appropriate statistics can be applied.

B. (U) <u>PROJECT TYPE:</u>

(S/NF) The various projects worked on by this office include foreign personalities, military related targets, event predictions, as well as search projects involving location of target personalities or moving equipment. Evaluation procedures with regard to search are very clear cut because either the location is accurate ("a hit") or it is not ("a miss"). Therefore, search information can be evaluated separately from broad categorical data. Training results are easier to evaluate

Approved For Release 2000/08/08 : CIA-RDP96-00789R002100030001-5

in terms of data accuracy since targets can be easily controlled or defined.

C. (U) <u>ROLE OF RECORDS/PROTOCOLS/PROCEDURES:</u>

(S/NF) The role of session records is an extremely important one. These records would include the people involved, information provided to the project personnel, project timing and other relevant data. Such project details are recorded and maintained in a permanent file or automated data bases. Specific protocols are also followed to insure proper records and other procedures are followed. A companion report, item 6 in the bibliography, contains protocol and methodology details.

(C) To further assist and improve the over-all evaluation process, future projects will be evaluated and assessed according to the procedures illustrated on figure 4. This flow diagram contains all the essential steps necessary for insuring that appropriate actions occur and range from task initiation through final data assessment and feedback. Details will be developed to clarify the various roles of each major phase and to identify guidelines for establishing uniform evaluation criteria.

Approved For Release 2000/08/08 : CIA-RDP96-00789R002100030001-5

DATA DEVELOPMENT, EVALUATION, AND ASSESSMENT PROCESS

Approved For Release 2000/08/08 : CIA-RDP96-00789R002100030001-5

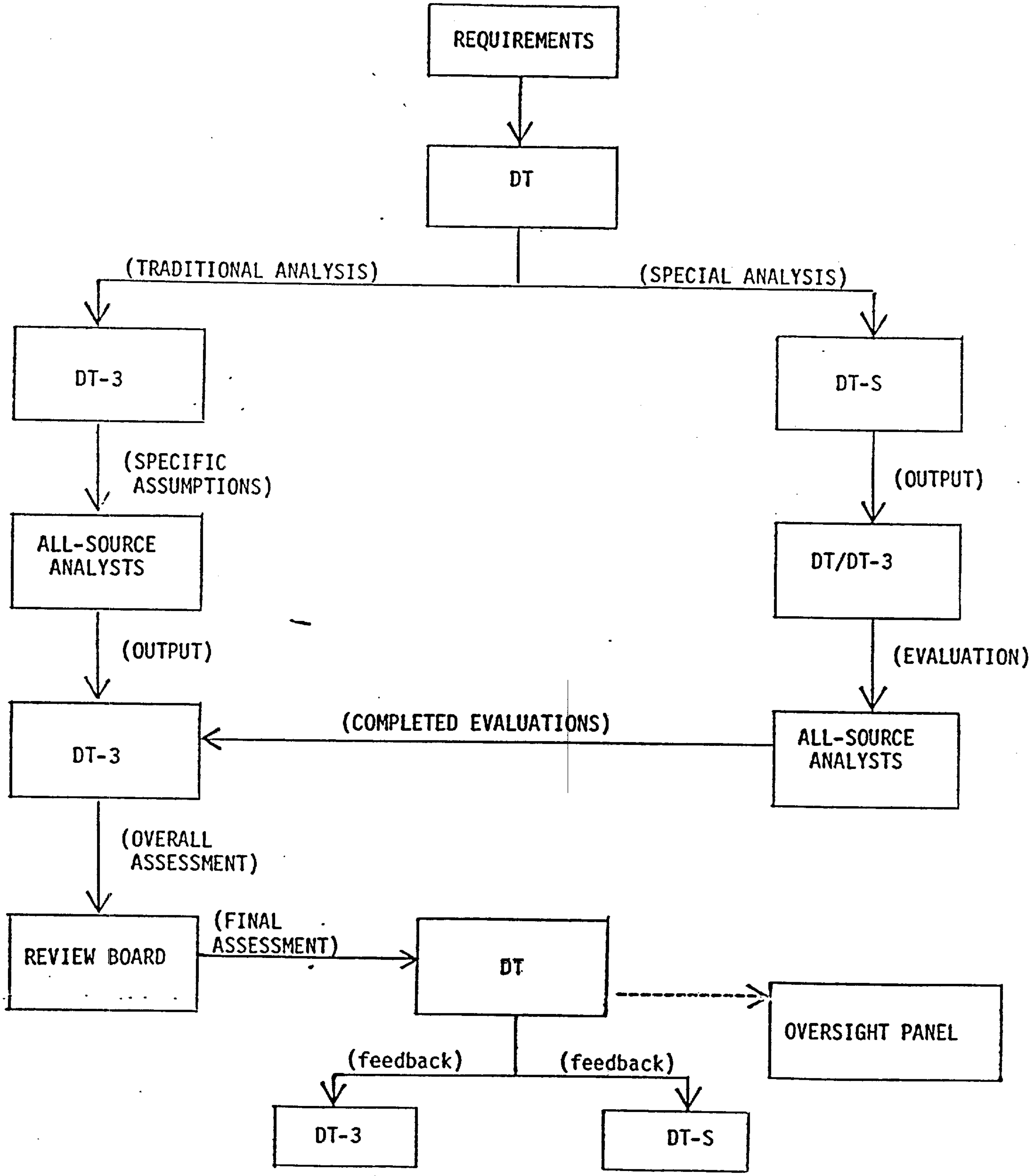


FIGURE 4

Approved For Release 2000/08/08 : CIA-RDP96-00789R002100030001-5

Approved For Release 2000/08/08 : CIA-RDP96-00789R002100030001-5

BIBLIOGRAPHY

1. "Sun Streak: Grill Flame Project Report (U)," Defense Intelligence Agency, Washington DC, (October 19, 1983), SECRET-NOFORN-WNINTEL-SAP.

2. "Special Report: Sun Streak Evaluation (U)," Defense Intelligence Agency, Washington DC, (November 2 1989), SECRET-NOFORN-WNINTEL-SAP.

3. Daniel Druckman and John A. Swets, Editors, "Enhancing Human Performance Issues, Theories, and Techniques" National Academy Press, Washington DC (1988), UNCLASSIFIED.

4. John A. Palmer, Charles Honorton and Jessica Utts," Reply to the National Research Council Study on Parapsychology", Parapsychology Association, Inc., Research Triangle Park, NC (1988), UNCLASSIFIED.

5. "Report of a Workshop on Experimental Parapsychology", International Security and Commerce Program, Office of Technology and Assessment, United States Congress (February 22, 1989), UNCLASSIFIED.

Approved For Release 2000/08/08 : CIA-RDP96-00789R002100030001-5

Approved For Release 2000/08/08 : CIA-RDP96-00789R002100030001-5

[redacted]

6. "Operational Project Protocols (U)," DIA Document No. DT-S-1006-SS, Defense Intelligence Agency, Washington, DC (November 27, 1990), SECRET-NOFORN-WNINTEL.

7. "SRI Studies in Remote Viewing (U)," Science Panel Report (March 1, 1984), SECRET-WNINTEL.

8. "Data Evaluation Issues," Paper, (Pending publication by DT-S), UNCLASSIFIED.

9. Beverly S. Humphrey et al., "Fuzzy Set Applications in Remote Viewing Analysis", Final Report--Objective A, Task 3 Covering the period 1 October 1986 to 30 September 1987, Project 1291, SRI International, Menlo Park, CA (December 1987), UNCLASSIFIED.

10. Edwin C. May, et al., "Applications of Fuzzy Sets to Rote Viewing Analysis (U)", Final Report--Objective F, Task 1 Covering the Period 1 October 1987 to 30 September 1988, Project 1291, SRI International, Menlo Park, CA (December 1988), SECRET-NOFORN-WNINTEL.

11. Edwin C. May, "An Application Oriented Remote Viewing Experiment (U)", Final Report Covering the Period 1 May 1988 to April 1989, Project 2740, SRI International, Menlo Park, CA

[redacted]

Approved For Release 2000/08/08 : CIA-RDP96-00789R002100030001-5

(April 1989), SECRET-NOFORN.

12. Wanda L. W. Luke et al., "A Prototype Analysis System For Special Remote Viewing Tasks (U)", Final Report--Task 6.0.3 Covering the Period 1 October 1988 to 30 September 1989, Project 1291, SRI International, Menlo Park, CA (October 1989), SECRET-NOFORN-WNINTEL.

13. Beverly S. Humphrey et al., "Remote Viewing Evaluation Techniques (U),", Final Report--Objective A, Task 4, Project 1291, SRI International, Menlo Park, CA (December 1986), SECRET-NOFORN-WNINTEL.

14. K. Ramakrishna Rao "The Basic Experiments in Parapsychology," McFarland & Company, Inc., Jefferson, NC, and London (1984), UNCLASSIFIED.