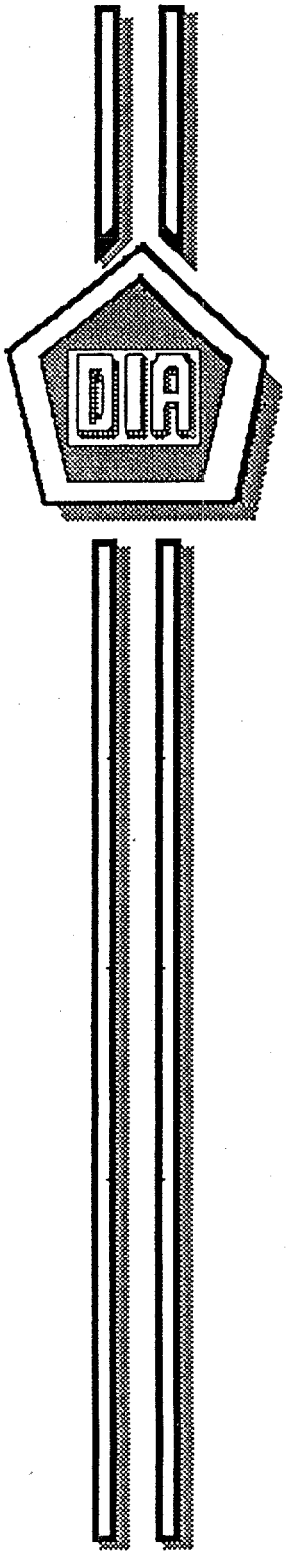
~~SECRET~~

DT-S-1046-SL



DEFENSE
INTELLIGENCE
AGENCY

PROFICIENCY PROGRAM: REPORTING AND EVALUATION (U)

3 JUNE 1992

NOFORN
~~SECRET~~
LIMDIS

STAR GATE

PROFICIENCY PROGRAM: REPORTING AND EVALUATION

SHORT TITLE: DTI-S-1046-SL

Date of Publication
3 JUNE 1992

This document was prepared by the
Directorate for Scientific and Technical Intelligence,
Defense Intelligence Agency.

Approved For Release 2000/08/08 : CIA-RDP96-00789R002900490001-7

TABLE OF CONTENTS

~~SECRET~~

PROFICIENCY PROGRAM; REPORTING AND EVALUATION (U)

1. (U) PURPOSE: The purpose of this memo is to outline an approach for reporting and evaluating the DT-S proficiency program.

2. (U) BACKGROUND:

(S/NF/SG/LIMDIS) The DT-S proficiency program was described in report DT-S-1039-SL, 21 June 1991, "Proficiency Enhancement Projects - Basic Approach". These projects are intended to aid in applications research and in performance trend assessments. They provide a systematic approach for addressing a variety of issues important to research and potential applications. In most cases, these projects provide rapid feedback that is essential for learning and for maintaining or improving skill levels. They also provide a means for viewers to explore new approaches and techniques; such exploratory activity could lead to new insights and data quality improvements.

(S/NF/SG/LIMDIS) The various types or categories of proficiency projects are: (1) exploration/calibration; (2) computer-aided; (3) beacon person; (4) geographic areas/sites; (5) personalities; (6) documents; (7) search/track; (8) precognition; (9) operational simulation projects (e.g., police cases, known facilities); and (10) research evaluation.

(S/NF/SG/LIMDIS) At present, target pools have been developed for several of these project types; others will be completed early in FY 1992. Some of these projects involve targets of opportunity that are scheduled only when certain situations occur, such as a search/track of a missing person. The wide variety of targets resulting from these categories helps maintain interest and poses new challenges.

3. (U) REPORTING:

(S/NF/SG/LIMDIS) Results of each proficiency project will be summarized in a separate summary report. This report will include a summary of project tasking/objectives, a description of the target material (or photo, when applicable), viewer summary data, and a table with an overall assessment/ evaluation of each viewer's results. Other material, such as trends/observations, may also be included in each summary report.

(U) These publications will be made available within two-three duty days following completion. Publication timing for each report may vary, however, depending on evaluation steps that are not yet finalized. Proficiency projects already performed to date will be summarized and published by 31 Nov 91.

(U) For most of the project categories, an additional summary report will be published. It will include data evaluation summaries for every group of ten projects to permit quick overviews of project results. Reporting for other proficiency projects, such as the computer-aided activity, will be based on a fixed time period such as every one to two months.

4. (U) EVALUATION:

(S/NF/SG/LIMDIS) The specific approach for evaluating proficiency enhancement projects depends on project category. For example, the computer-aided project involves a forced choice (one out of four) procedure. Statistical data can easily be developed and probabilities can be computed. A dedicated computer program for calculating appropriate probabilities has been developed and is in use.

(S/NF/SG/LIMDIS) Other proficiency projects require the use of target pools (e.g., beacon person, photos of geographic areas) that can only be evaluated by subjective/objective approaches. Several evaluation approaches will be used for such projects, and are based on methods developed by the research community. These methods are scale value assessment, rank-ordering, and a detailed assessment of all data elements. This detailed assessment is based on a modified "fuzzy-set" analysis methodology originally developed for artificial intelligence (AI) applications. (Additional background is in DT-S publication, DT-S-1010-S, 13 Dec 1990, "Evaluation Methods").

(S/NF/SG/LIMDIS) The scale value assessment method follows from an earlier approach that proved to be expedient and reasonably accurate. Specifically, evaluation forms similar to those on attachment 1 will be completed as soon as possible after each proficiency project. Data will be generally evaluated for key elements in two basic data types: (1) concept/generic; and (2) analytic labeling. The scale values (0-5) will be generally as defined on attachment 2, although other proficiency categories may require different specifics. Evaluators will compare the viewer's data to the actual target and make judgements as to degree of target/data correlation according to the scale criteria.

(S/NF/SG/LIMDIS) Evaluators will need to have a uniform understanding of evaluation perspectives. Attachment 3 illustrates a first-cut approach to evaluation guidance development. A memo with detailed evaluation guidance will be developed by 31 Nov 1991. To facilitate evaluation, viewers will also need to prepare summaries in a consistent style. Appropriate guidance will be developed and made available for the viewers.

SG1J (S/NF/SG/LIMDIS) The evaluators will consist of two individuals from within DT-S, and a third individual from outside of DT-S.
SG1J The DT-S individuals will be [redacted] (due to his prior experience in working with such data), and [redacted] (a newcomer with good objectivity). The individual outside of DT-S needs to be identified; however, this can be any analyst who is readily available and is reasonably neutral toward this area.

(S/NF/SG/LIMDIS) A rank-ordering method will also be occasionally used. In this method, judges/evaluators compare and rank the viewer's data to four (or more) possible targets, one of which is the correct target. This method is amenable to statistical evaluation; however, it is time consuming and requires a large "dummy" target pool. This process generates interest and excitement for the viewers since they can also serve as one of the judges and take an active role in the evaluation process.

(S/NF/SG/LIMDIS) The quantified evaluation approach using fuzzy set techniques will also be applied to evaluating some proficiency projects. This method is very time consuming, although it can be simplified and can eventually become a standard evaluation approach along with the other simpler methods. A modified fuzzy set analysis procedure should be in-hand within three months for some of the proficiency target categories. It is probably not time efficient to apply this method until remote viewing trends indicate improvement in overall data quality. This method would be appropriate for complex proficiency projects, such as operational simulations.

(S/NF/SG/LIMDIS) Involvement in the upcoming R/D projects (now in final planning) will greatly assist us in improving our knowledge and use of the evaluation techniques planned for our proficiency projects. The R/D projects will be evaluated on rank-ordering, scale value assessment, and a comprehensive fuzzy set approach.

PROJECT NO. _____

EVALUATION RECORDS

PROFICIENCY PROJECTS

Source	Evaluation Categories (For key elements)	Proficiency Coordinator (DT-S)	Analysis Specialist (DT-S)	Outside Reviewer ( )	Other
018	a. Concept/Generic				
	b. Analytic labeling				
025	a. Concept/Generic				
	b. Analytic labeling				
049	a. Concept/Generic				
	b. Analytic labeling				
052	a. Concept/Generic				
	b. Analytic labeling				
079	a. Concept/Generic				
	b. Analytic labeling				
____	a. Concept/Generic				
	b. Analytic labeling				
____	a. Concept/Generic				
	b. Analytic labeling				
____	a. Concept/Generic				
	b. Analytic labeling				
____	a. Concept/Generic				
	b. Analytic labeling				

ATTACHMENT 1

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
2	30-50	
3	50-70	MIXTURE OF RELEVANT AND INCORRECT DATA WITH MAJORITY CORRECT AND UNAMBIGUOUS
4	70-90	
5	90-100	VERY HIGH TARGET CORRELATION WITH ESSENTIALLY NO AMBIGUOUS DATA

GUIDANCE FOR EVALUATORS

The data in this report were obtained through a unique methodology. Some background and qualifiers regarding data peculiarities may be useful as evaluation guides. Generally, the information obtained through this technique is likely to consist of a mixture of correct, incorrect or irrelevant data. In addition, certain types of data categories are usually more reliable than others.

Conceptual descriptive data tends to be more reliable than analytic labeling. As an example, a source may inappropriately report (label) a body of water at a designated target as a "recreational swimming pool," when in fact it is a water purification pool or a fishing pond. As another example, an aircraft fuselage may be erroneously labeled by the source as a submarine hull. Caution is therefore advised when a source's descriptions evidence a great deal of analytic labeling. Other anomalies may exist in the overall descriptive reporting. Certain perceptions, such as quantities, may be highly problematic. A target having a multi-building complex may, for example, only be perceived as having one or two buildings.

There are other data peculiarities, such as difficulties in scale or spatial perception. These aspects that will be described in detail through appropriate briefings.

ATTACHMENT 3