

July 16, 2008

Mr. Philippe Jamet, Director  
Division of Nuclear Installation Safety  
Department of Nuclear Safety and Security  
International Atomic Energy Agency  
Wagramer Strasse 5, PO Box 100  
1400 Vienna, Austria

Dear Mr. Jamet,

This is to follow-up on our discussions on May 9, 2008, and to my May 30, 2008 letter to you in which I stated that I would send NRC comments on the February 2008 version of the IAEA's Safety Culture Assessment Review Team (SCART) Guidelines. I appreciate your invitation to provide comments to you now versus waiting until another draft of the SCART Guidelines is developed.

Members of the U.S. Nuclear Regulatory Commission's (NRC's) Safety Culture Working Group (SCWG) reviewed the June 2007 draft of the SCART Guidelines and I provided the Working Group's comments to the IAEA. We also reviewed the November 2007 draft as well as the February 2008 final version of the Guidelines. Based on our review of the latest version, the SCWG members have continuing concerns about the SCART Guidelines methodology. The enclosure provides our comments on the February 2008 Guidelines. In general, the comments relate to the Guidelines methodology providing insufficient data to reach reliable conclusions on a facility's safety culture (i.e., insufficient data derived from the SCART assessment methods, i.e., interviews, document reviews, and observations, and only assessing one safety culture characteristic at a time versus a more holistic approach). If there is need for clarification on the comments, we would be happy to discuss them with you and respond to any questions you may have.

SCWG members are concerned that weaknesses in the SCART Guidelines methodology may result in the SCART not adequately identifying safety culture issues at a facility. This could result in a facility believing that its safety culture is generally healthy, without awareness of weaknesses that need to be addressed, thus impacting safety performance. Furthermore, certain facilities and regulatory bodies could adopt or adapt the SCART Guidelines for their own use, not being aware of the weaknesses in the assessment methodology.

Because of the influence the SCART mission can have on a facility's safety performance we offer the following suggestions. In the near term, the IAEA should consider all the comments they have received on the SCART Guidelines from NRC and other commenters. Following IAEA review, the IAEA should consider providing the resolution of the comments to the submitters in the interest of transparent decision-making and so that the submitters understand the basis for the decisions to accept or reject comments. In addition, as I suggested in my May 30<sup>th</sup> letter, we recommend that IAEA sponsor a subject matter expert workshop on "Criteria for Evaluating Safety Culture Assessment Methodologies," at the earliest opportunity and revise the SCART methodology based on the results of the workshop.

In closing, the IAEA has long been at the forefront of promoting the importance of safety culture of nuclear facilities since the term “safety culture” first appeared in the International Nuclear Safety Advisory Group (INSAG) “Summary Report on the Post-Accident Review Meeting on the Chernobyl Accident,” published by IAEA as INSAG-1 in 1986. INSAG further defined safety culture in INSAG-4, published by IAEA in 1988. Subsequent INSAG and IAEA publications further refined the description of the characteristics and attributes important to a healthy safety culture. More recently IAEA incorporated safety culture guidance in the IAEA safety standards, e.g. in GS-R-3 “The Management System for Facilities and Activities” and GS-G3.1 “Application of the Management System for Facilities and Activities.” Because of its leadership role in promoting safety culture, it is important that IAEA apply generally accepted safety culture assessment criteria to its own methodology.

Again, thank you for the opportunity to provide comments and I look forward to further communications on these topics.

Sincerely yours,

/RA/

Isabelle Schoenfeld  
Senior Program Manager and  
Chair, NRC Safety Culture Working Group  
Office of Enforcement  
U.S. Nuclear Regulatory Commission

Enclosure: As stated

In closing, the IAEA has long been at the forefront of promoting the importance of safety culture of nuclear facilities since the term “safety culture” first appeared in the International Nuclear Safety Advisory Group (INSAG) “Summary Report on the Post-Accident Review Meeting on the Chernobyl Accident,” published by IAEA as INSAG-1 in 1986. INSAG further defined safety culture in INSAG-4, published by IAEA in 1988. Subsequent INSAG and IAEA publications further refined the description of the characteristics and attributes important to a healthy safety culture. More recently IAEA incorporated safety culture guidance in the IAEA safety standards, e.g. in GS-R-3 “The Management System for Facilities and Activities” and GS-G3.1 “Application of the Management System for Facilities and Activities.” Because of its leadership role in promoting safety culture, it is important that IAEA apply generally accepted safety culture assessment criteria to its own methodology.

Again, thank you for the opportunity to provide comments and I look forward to further communications on these topics.

Sincerely yours,

/RA/  
 Isabelle Schoenfeld  
 Senior Program Manager and  
 Chair, NRC Safety Culture Working Group  
 Office of Enforcement  
 U.S. Nuclear Regulatory Commission

Enclosure: As stated

- |                           |                     |
|---------------------------|---------------------|
| cc: C. Viktorsson, IAEA   | R. Clark, IAEA      |
| K. Dahlgren-Persson, IAEA | C. Carpenter, OE    |
| S. Magruder, OE           | D. Solorio, OE      |
| V. Barnes, RES            | M. Virgilio, DEDMRT |
| D. Diaz-Toro, OEDO        | H. Astwood, OIP     |
| J. Schwartzman, OIP       | S. Dembek, OIP      |
| S. Moore, OIP             |                     |

**ML081700408**

OFFICE	OE	OE	OIP
NAME	I. Schoenfeld	S. Magruder for C. Carpenter	S. Moore for M. Doane
DATE	06/18 /08	07/16/08	7/16/08

OFFICIAL RECORD COPY

## U.S. NUCLEAR REGULATORY COMMISSION COMMENTS ON THE IAEA SAFETY CULTURE ASSESSMENT REVIEW TEAM (SCART) GUIDELINES

The comments are presented in the order of the major sections in the SCART guidelines. The words in italics are quotations from the guidelines.

### Introduction

*“As safety culture relates to attitudes, behavior norms and human values, the SCART guidelines offer a method, which is based on behavioral science, and is transparent enough to be widely understood and accepted.”*

1. It is unclear how the assessment methodology reflects behavioral science. Fundamental principles of behavioral science require that assessment methodologies are standardized, reliable and valid. Although the guidelines describe a standardized process for rating safety culture characteristics and resolving intra-team differences of opinions, they should emphasize standardization in how information about the organization is collected; it appears that the amount of information to be collected should be increased or expanded to support reliable conclusions.

*“The reviewers are either peers from similar nuclear organizations in other member states or behavioral consultants with direct experience in assessment and enhancement of safety culture in nuclear facilities.”*

2. The above quote indicates that including behavioral consultants as team members is an option. However, pg. 20 states that “Behavioral scientists (at least two reviewers) as well as technical specialists with a solid grasp of behavioral scientific concepts should be part of the team.” We agree with the latter to support a valid assessment. Therefore, that behavioral scientists should be included on the team should be clarified on pg. 5 of the introduction.

### SCART Review Essentials

*“SCART does not analyse the management system and is not a substitute for a comprehensive review of the management system...”*

3. The IAEA requirements documents on integrated management systems (MS) clearly state that safety culture is cross-cutting in the MS. To that extent, the SCART guidelines should include a review of whether and how the MS addresses safety culture if the facility implements an integrated management system.

*“Members of Bd. of Directors and CNOs and Executive Officers are in most cases corporate functions. They are included in the SCART review only as far as their activities relate to the safety culture of the nuclear facility.”*

4. The interview schedule in Table 1 of the guidelines should include corporate management. In order to provide reliable insights, safety culture assessments must be comprehensive. Therefore, the assessment should address all levels of the organization, including the CNO and other corporate-level managers who are linked to the facility.

Because the span of control and influence of individuals at this organizational level is so far-reaching, their attitudes, values and expectations typically set the stage for the health of the entire organization's safety culture. Interviewing these individuals is necessary to obtain a full picture of their influence on the safety culture, as well as to determine the extent to which persons at lower levels in the organization are aligned with the corporate levels' attitudes, values and expectations.

5. Throughout this section, the SCART refers to SCART reviewer "opinions." This section should acknowledge that collecting a sufficient amount of data through the application of effective, reliable assessment methods/tools, reduces the reliance on reviewer "opinions." Sufficient, information provides the basis for reviewers to analyse the data to reach reliable conclusions.

### SCART Review Method

*"The nuclear facility should be analyzed as a whole."*

#### Interviews

6. This statement suggests that the assessment should be comprehensive, which implies that the SCART should assess all levels of the organization as well as all functional departments that have an impact on safety. However, the guidelines indicate that not all of the functional areas of the organization that have an impact on safety are assessed. Further, and significantly, the guidelines require that only frontline (shop floor) personnel in operations and maintenance departments be interviewed. Frontline personnel from other functional groups who have a clear impact on safety, such as engineering and health and safety, are not to be interviewed, although the supervisors/mid-and senior management of these departments are interviewed. This is a significant weakness because it's not only important to understand the degree of alignment among the frontline/supervisors/mid-and senior management in a safety culture assessment, as discussed above, but also because the frontline is the group in the organization that "touches" the equipment. Therefore, frontline attitudes and behaviors are likely to be the most valid source of normative information about "how we really do things around here." Inadequate sampling of frontline personnel will provide an incomplete or inaccurate picture of what's going on in the organization. The guidelines provide no basis for limiting the scope of frontline personnel interviewed to operations and maintenance. In addition, it would be challenging to be able to assess inter-departmental communication and coordination if the assessment omits interviewing key groups who are responsible for important aspects of the workflow.

7. Table 1 provides an example schedule which shows that of 88 interviews scheduled, 8 are with senior managers, 45 are with middle managers, 25 are with first line supervisors, and 10 are with the shop floor. Hence the schedule is heavily biased towards manager/supervisor interviews.

8. Another constraint is that each of the interviews is "kept within a 60 minute time span." This is a very short time period considering that there needs to be time for the interviewer to provide introductory remarks and to respond to any questions that the

interviewee might have in the beginning or the conclusion of the interview. If a translator is needed, this also limits the time to actually gain information and to ask follow-up questions, if needed.

9. The SCART guidelines should provide guidance on the interview sampling methodology. A sampling methodology would help ensure comprehensive coverage of all functional areas and all organizational levels. Without a sampling methodology there is little assurance that the SCART is interviewing a sufficient number of people and type of positions to be able to identify the outliers and to reach reliable conclusions.

10. The SCART guidelines provide a list of suggested themes and guiding interview questions in Annex II for each safety culture attribute. The guidelines should clarify how the questions relate to the attribute. Further, information should be provided on how the suggested questions would be used in the different SCART missions, i.e., since they are “suggested,” what guidance should be followed to modify them for a particular SCART mission.

### Scoring System

*The SCART process, has two essential features that give it checks and balances and protect it from the biases of the team members: (1) individual evaluations with qualitative and quantitative output (comments and scoring system), (2) managed, in-depth team discussion which arises from the individual comments and ratings.”*

11. The scoring system is one method to attempt to impose “checks and balances” on team member biases, but it is only one method at one point in the process, and therefore only partially effective. There are other opportunities for biases to creep into an assessment, and the process should include “checks and balances” to minimize the potential effects of biases at each step where they might occur. These “checks and balances,” include, for example, training and practice for each team member in conducting interviews in a manner that does not inadvertently influence the interviewees’ responses; standardizing and practicing the conduct of behavioral observations to ensure that the observers’ evaluations are consistent and reliable; and conducting a sufficient number of interviews and observations to ensure that the information derived is representative, rather than an unusual circumstance. Individual interview styles, “tunnel vision,” inadequate sample sizes, and random influences resulting from who the team happened to contact that day create openings for many types of biases to invalidate the team’s conclusions.

12. The scoring and team discussion step should be addressed earlier in the assessment of each safety culture characteristic. If there are differences of opinion between team members, the guidelines should provide for the opportunity to collect additional information focused on resolving the discrepancy. Although there will likely always be occasions when team members will disagree no matter how much empirical information is collected, gathering focused follow-up information often permits a team to reach a data-based, rather than an opinion-based, consensus.

## Pre- SCART Mission Process

*The nuclear power plant develops and applies a staff questionnaire survey as a form of self-assessment of safety culture.” “To minimize low response rates and biases, the staff questionnaire should be prepared and executed with professionalism and utmost care... The data from the survey should be available to the SCART team at least one month before the SCART mission.”*

13. The SCART guidelines discuss the use of a self-assessment safety culture questionnaire/survey by the facility as an option, with the decision to be made between the SCART leader and the installation’s representatives, rather than as a standard expectation. A well-designed and administered survey is one means to obtain a standardized and wide-scope assessment of workforce attitudes and perceptions of the safety culture.

14. The guidelines should indicate how the decision to administer a survey will be reached. The guidelines should also provide safety culture self-assessment criteria to assist the facility in developing and administering the survey. The guidance “*to minimize such biases, the staff questionnaire should be prepared and executed with professionalism and utmost care*” is not sufficiently detailed to assure that the survey will meet acceptable survey criteria. Further, the behavioral scientists on the SCART should review the survey before it is administered to determine its adequacy, because the team would use the data from the survey to inform their assessment. If the survey itself is inadequate, if it is administered in a manner that could influence how individuals respond to it, or if only a small percentage of the facility’s population participate in the survey, then the survey results will be limited. Further, if the SCART has not evaluated the survey method’s adequacy, the survey results could provide misleading information.

## SCART Team Composition

*The deputy team leader is an experienced outside safety culture expert or an IAEA staff member.*

15. The guidelines acknowledge that a deputy team leader should be a “safety culture expert.” However, it appears to limit the designation of the safety culture expert to the deputy team leader’s position only in circumstances where the position is filled by an outsider. This suggests that if the deputy position is filled by an IAEA staff member, that staff member does not need to be a safety culture expert. No rationale is given for this distinction. We suggest that one of the leadership positions, either the deputy position or the team leader position should be filled by a safety culture expert whether they are from outside or within the IAEA.

## SCART Mission

### SCART Mission Schedule: Arrangements for interviews and observations

The SCART guidelines appropriately use the safety culture characteristics and attributes developed by the IAEA. There are five characteristics and 37 attributes distributed within the five characteristics. The five characteristics are:

Characteristic A: Safety is a clearly recognized value

Characteristic B: Leadership for safety is clear

Characteristic C: Accountability for safety is clear

Characteristic D: Safety is integrated into all activities

Characteristic E: Safety is learning driven

16. The SCART approach has the team evaluating only one safety culture characteristic/day. This approach is conceptually inconsistent with the definition of safety culture, in that each characteristic is treated separately rather than integrated with information about all of the other characteristics, to arrive at insights about the installation's safety culture as a whole. Safety culture comprises more than the "sum of its parts" (i.e., the characteristics and attributes overlap and are interrelated but are neither mutually exclusive nor exhaustive, and cannot be summed). To collect data on only one characteristic per day creates an artificial separation among the characteristics and severely limits the data's usefulness.

17. This approach seriously limits the amount of information (sample sizes) that will be collected on any individual characteristic, which is one of the greatest weaknesses in the guidelines. According to the guidelines, for each safety culture characteristic an average of 16 interviews will be conducted. If there are four team members and they individually conduct the interviews, this means that each would base his or her individual rankings and opinions on approximately four interviews per characteristic. This is a very limited amount of information and unlikely to support reliable conclusions.

The SCART should make maximum use of each information-collection activity (i.e., interviews, observations, document reviews) to gather data related to all of the safety culture characteristics with each use vs. one/day. The team's conclusions would then be based on the whole of the information collected.

### SCART Mission Schedule: Interview/Observation

18. The following considerations also limit the usefulness of the approach in ensuring that the team reaches reliable and valid conclusions.

- a) Formulation of recommendations, suggestions, and good practices occurs after the first week for Characteristics A, B, and C and after the end of the second week for Characteristics D, and E. As previously discussed, this approach reflects a lack of appreciation for the interconnectedness of the safety culture characteristics and will not yield results that address their relatedness or the whole of the safety culture.

- b) Senior management interviews should be conducted throughout the assessment, not just on the last day. These interviews should occur before the team has formulated recommendations, suggestions, and good practices for safety culture characteristics A, B, and C so that any information gained from interviews with senior managers (e.g., how their views align with the views of the other organizational levels on these characteristics) would be considered in this formulation. Senior managers have a strong influence on the organization's safety culture and can support or undermine a safety culture (e.g., at a minimum by their attitudes and behaviors).
- c) It appears that there is one interviewer for each interview. For greater reliability it is very useful to have two interviewers for each interview (a behavioral scientist and an engineer/nuclear technical specialist for each interview). This would also provide additional insights from each professional perspective. In addition, two-person interview teams allows for one person to take notes while the other person conducts the interview.

### SCART Attribute Evaluation

19. The guidelines state *“that when rating the attributes, the reviewers may be supported by behavioral descriptions (anchors) for each attribute.”* These behavioral descriptions are not defined but state that they *should be specifically tailored before the SCART mission and they should consider the national or regional culture of the area where the nuclear facility is located.* The guidelines should include supporting information on why this should be considered and information on how it will be done. It would be quite a challenge to understand the “culture” of a nation or region and how it relates to the culture of the facility and develop behavioral descriptions related to it. Any assessment in this area would need to be supported by sociological peer-reviewed research.

### Miscellaneous

20. There should be additional guidance on developing the data collection tools (e.g., interview protocols, observation check-lists) and training in applying the methods. These are fundamental to mitigating potential team member biases and “tunnel vision.” If there are basic flaws in what one of the team members is asking about (content) or how the individual is doing the data collection, it will be too late to correct the problem(s) by the time the team does their ratings and begins team discussions.

21. The guidelines dedicate few team resources for observations. Most observations are done only once for any activity over the two-week period. It is unlikely that one observation of an activity can provide valid and reliable information. In addition, it appears that one person conducts each observation; two observers would provide greater reliability.

22. Document reviews should include event histories, root cause analyses, or other documentation related to them such as how the facility developed and applied lessons learned or corrective actions. Without this information, the team would miss the opportunity to gain insights into Characteristic E.