



UNITED STATES
NUCLEAR REGULATORY COMMISSION

WASHINGTON, D.C. 20555-0001

June 23, 2005

MEMORANDUM TO: Roy P. Zimmerman, Director
Office of Nuclear Security and Incident Response

FROM: Bruce A. Boger, Director *BABoger*
Division of Inspection Program Management
Office of Nuclear Reactor Regulation

SUBJECT: PANEL'S REVIEW OF THE DIFFERING PROFESSIONAL OPINION
INVOLVING FORCE-ON-FORCE EVALUATION CRITERIA

In response to your memo dated January 24, 2005, the ad hoc review panel (the "Panel") has reviewed the subject Differing Professional Opinion (DPO). A statement of the submitter's concern, background information, and the Panel's review process, conclusions and recommendations are provided below.

Background

In a memorandum dated January 24, 2005, you appointed me as Chairman of the ad hoc review panel for the DPO involving Force-on-Force (FOF) evaluation criteria. Jack Davis and Dennis Vernon, both from NSIR/DNS, were also designated by you as Panel members. Rani Franovich of NRR/DIPM assisted the Panel.

Statement of Concern

The Panel reviewed the written concerns submitted in the DPO (Attachment 1) and met with the submitter on January 27, 2005. Following introductions, a discussion of the NRC's DPO program, and a description of the process, the Panel asked the submitter to summarize his concerns. The Panel met with the submitter again on February 9, 2005, to obtain clarification on certain details of his concerns.

The Panel provided a concise summary of the concern to the submitter for his review and comment on February 21, 2005. The summary was discussed with the submitter on several occasions and revised based upon his comments. A final summary was approved by the submitter on April 11, 2005, and is as follows:

In short, the submitter contends that a win/lose paradigm is employed by the NRC in evaluating Force-on-Force (FOF) exercises at power reactor facilities. The submitter also contends that the win/lose approach is insufficient in that:

- (1) it does not adequately address relevant insights into licensee performance during the conduct of FOF exercises (i.e., is not performance-based);

CONTACT: Rani L. Franovich, NRR/DIPM
301-415-1868

(2) results derived from these win/lose criteria are not reliable and do not provide defensible results and actions taken, or not taken, by the NRC and respective licensees due to these results may be detrimental to the public health and safety.

(3) there are numerous artificialities, many of which are complex and significant, associated with these exercises and these artificialities routinely affect the outcome of exercises. Most of the consequences resulting from these artificialities are known and can be adequately factored into the evaluation process. However, there are consequences created by some artificialities that evaluators are either unaware of or do not fully understand and these consequences can be the determining factor in which side wins or loses an exercise

(4) the affect of drill artificialities on the final outcome (success or failure to protect the target set) of the FOF exercises is amplified in the win/lose paradigm.

Regardless of the criteria used artificialities will always affect the outcome of these types of exercises. The number and type of artificialities, and the subsequent consequences relating to the bottom-line, remain the same regardless of the evaluation criteria used. However, a performance based criteria that evaluates the capabilities of protective force attributes such as Command and Control, Communications, Tactics (individual and team), Equipment, Protection Strategy, etc. along with other supporting protective systems is not reliant on the bottom-line results to obtain necessary and beneficial data. Using performance based criteria evaluators are able to determine the effectiveness of these individual attributes regardless of who wins or loses the exercise. Evaluating these essential attributes provide evaluators with information pertaining to the protective force/systems' effectiveness which in turn provides a basis for determining whether the licensee has the necessary capabilities to adequately protect essential equipment.

(5) It precludes adequate evaluations of "wins" to identify performance deficiencies and areas for improvement;

(6) It provides no incentive to licensees to function in an open, legitimate manner and may provide motivation to manipulate their performance in an effort to "win."

Submitter's Proposal

A final summary of the submitter's proposal was approved by the submitter on April 11, 2005, and is as follows:

The submitter proposed that the NRC adopt a different evaluation approach similar to one employed by the Department of Energy. The proposed methodology, as the Panel understands it, is as follows:

Revise the FOF evaluation criteria. Employ performance-based criteria that focus on security force attributes (e.g., command and control, tactics, communications, etc.). An appropriate number of NRC evaluators with appropriate training should be used to assess protective force performance. Achievement of target set(s) provides insights into protective force performance, but the focus of performance assessment should be on the security force attributes and supporting protective

systems. The focus of assessment should not be on win/lose. Win/lose is important to the assessment but the focus should be on why the licensee won or lost with emphasis on performance-based criteria.

In summary, the submitter contends that when using the Win/Lose criteria a licensee with a less than adequate protective system/force can "win" exercises while an licensee with adequate protective measures can "lose" due the existence of artificialities. This is especially significant when evaluators are either unaware of the introduction of one or more artificialities or the potential consequences of these artificialities. Moreover, the perceptions and actions taken based on these indefensible results can be counterproductive to the security posture at the evaluated facility and detrimental to credibility of the NSIR Force-on-Force program.

Review Process

The objective of the Panel was to conduct a thorough review of the submitter's concerns and proposal and make a recommendation to you regarding the disposition of the issues presented therein. To accomplish this, the Panel sought to (1) verify that the submitter's understanding of the NRC's FOF evaluation criteria was correct; (2) to gain insights into NRC policy governing FOF exercises (e.g., how the evaluation program was developed, how it has evolved over time, and the future direction of the program); (3) determine the extent to which the submitter's concern had been expressed to and discussed with his management; and (4) evaluate the merits of the concern and proposal.

To this end, the Panel reviewed a number of NRC and DOE documents. The Panel also interviewed a number of NSIR/DNS staff and management, as well as representatives of DOE responsible for the DOE FOF program. A detailed presentation of the Panel's observations and conclusions based upon the information it reviewed is contained in the Panel's report (Attachment 2). A list of references is provided in Appendix 1 of that report.

As noted in the report, in the Panel's view, the submitter's concerns were not well-communicated up the NSIR management chain prior to the submittal of the DPO. Although opportunities existed, a thorough vetting of the submitter's concerns did not take place to ensure that they were fully considered and dispositioned by NSIR management. There was no documentation that could be used to review the management decision-making process on the concerns that were expressed by the submitter. The Panel had to rely on numerous interviews to develop its understanding of how management addressed the issues and how the security staff implemented the new FOF program. Furthermore, the rapidly evolving nature of the FOF program over the previous two years and the absence of a clearly articulated future vision for the program revealed inconsistencies between information provided in a written format and program descriptions provided by staff and management. As a result, the Panel received necessary information related to the DPO and some additional information that was related to the FOF program in general. This allowed the Panel to develop conclusions and recommendations specifically related to the DPO, and to also develop insights related to the FOF program in general, which are offered for consideration.

Panel's Conclusions

The Panel recognizes that the FOF exercise is just one element that is considered when the effectiveness of a licensee's physical security program is evaluated. However, FOF exercises

are the only means by which the NRC can observe and evaluate a protective force's implementation of its defensive strategy. These exercises provide valuable insights into a protective force's ability to thwart adversaries that can not be obtained by reviewing a licensee's security plan or conducting the Security Baseline Inspection Program. Although NRC documentation does not refer to win/lose, the Panel agrees that the appearance of a win/lose paradigm exists due to the emphasis placed on target set protection in the significance determination process (SDP) along with the staff's understanding that protective force performance deficiencies exhibited during FOF exercises can only be developed as findings if they are linked to the loss of a target set. The NRC staff has been discouraged from bringing to the licensee's attention performance deficiencies exhibited when a target set has been protected due to a concern that this would be viewed as consulting with the licensee.

The submitter expressed in some detail his concerns regarding the influences of FOF exercise artificialities and the potential consequences on the outcome of an exercise. The Panel acknowledges that the influences of artificialities are well-known to the FOF program staff and that considerable efforts have been undertaken to reduce this influence. However, even when these artificialities are minimized, they can not be eliminated and they have the potential to invalidate or render indeterminate FOF exercise results. The submitter also asserts that in some cases artificialities exist that are unknown to the evaluators and may impact the protection of target sets. For these reasons, the Panel finds merit in the submitter's concerns with the potential impact of artificialities on FOF exercise outcomes.

The submitter contends that the FOF evaluation criteria should be revised to provide a focus on protective force attributes such as command and control, tactics, and communications. His proposal would retain consideration of drill outcomes in terms of protecting target sets; however, it would add to NRC evaluations a broader assessment of the protective force's ability to exhibit the performance attributes necessary to successfully execute a defensive strategy. There is some evidence that the staff currently applies an informal and team-specific approach to assessing performance attributes. However, the Panel feels that a structured assessment of performance attributes would enhance the NRC's ability to understand and articulate why a protective force was successful in its defense of target sets as well as enhance the NRC's ability to determine whether licensee corrective actions are appropriate in those instances where target sets were not well-protected. Such an assessment would also provide a template to consistently evaluate the adequacy of the licensee's critique of a FOF exercise. In the Panel's view, a structured assessment with established criteria would also provide the basis for discussions with licensees on performance deficiencies, regardless of target set outcomes and without the appearance of consulting. The Panel agrees with the submitter's concept.

The submitter proposed that the NRC adopt an evaluation approach similar to the one employed by the Department of Energy (DOE). The Panel found the DOE FOF program, with its 25 years of experience, to be mature and worthy of consideration. However, the Panel did not fully investigate the merits of migrating to an approach similar to one employed by DOE. Although the regulatory relationship that exists between the NRC and its licensees is much different than the relationship between DOE and its contractors, the Panel determined that improvements can be gleaned from this program to improve the existing NRC program. The resulting process would involve consideration of protective force performance across a number of performance attributes and the impact of artificialities on drill outcomes. Specifically, a hybrid program that retains aspects of the current FOF Program, augmented with features of the submitter's proposal or alternative changes that minimize the impact of artificialities on the

evaluation criteria applied to the protective force, could optimize the NRC's FOF performance assessment capabilities within budgetary constraints. A hybrid program would also allow the NRC to borrow evaluation concepts and strategies that are appropriate for the regulatory relationship NRC has with commercial power reactor facilities, which the Panel recognizes is quite different from the relationship between the DOE evaluating arm and the facilities it evaluates. It could also provide a vehicle, consistent with Reactor Oversight Process (ROP) principles, to identify and provide feedback to licensees on performance deficiencies.

The submitter also posits that the proper evaluation of protective force attributes would require specific training for NRC evaluators. He also asserts that the number of evaluators would have to be appropriate to adequately monitor a representative sample of attributes to compensate for artificialities. The Panel acknowledges the staff's efforts to recruit security experts and to develop a training and qualification process for NRC inspectors. However, it appears to the Panel that more structured and specialized training and periodic retraining and refreshing of perishable skills is likely to be needed for inspectors with varying levels of security experience to consistently perform the hybrid program noted in the preceding paragraph. The Panel notes that an assessment of the revised hybrid program, including the sample size of performance attributes, would have to be conducted to determine the appropriate number of evaluators.

The Panel concludes that the concept proposed by the submitter has merit and, if applied, would allow the NRC to more reliably assess the capability of a protective force to execute an effective defensive strategy. The Panel believes that a hybrid approach that retains the goal of protecting target sets and applies the performance assessment tools and techniques proposed by the submitter is viable, comports with the principles of the ROP, and would be more effective in both assessing and improving performance.

Panel's Recommendations

With respect to the NRC's evaluation of FOF exercises, the Panel recommends consideration of the following:

1. Benchmark with other federal agencies (DOD, DOE) to gather insights, techniques and strategies for evaluating protective force performance during FOF exercises.
2. Develop protective force performance attributes and evaluation criteria that should be used by FOF evaluators.
3. Incorporate protective force performance attributes and evaluation criteria into inspection procedures, assessment tools (including those for the SDP and licensee feedback), and guidance documents. Document assessments in inspection reports.
4. Develop a structured process to provide licensees FOF exercise feedback on protective force performance against the attributes and evaluation criteria identified.
5. Perform a job-task analysis for NRC FOF evaluators. Develop a more structured and specialized training and certification program (including refresher training) for all FOF evaluators.

6. **Employ a Communication Plan to reach out to internal and external stakeholders to develop a consistent understanding of the revised program.**

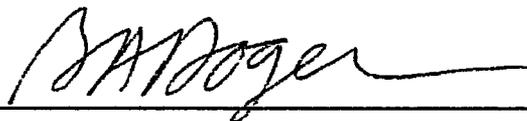
With respect to the Panel's general observations of the security cornerstone and oversight process, the Panel recommends consideration of the following:

1. **Revisit the FOF program's objectives. Clearly and consistently articulate those objectives in program documents, procedures, and communications with internal and external stakeholders.**
2. **Ensure that FOF program documents are complete and provide guidance to ensure that performance issues are identified, developed into findings, and documented in inspection reports. This will enable the staff to trend performance over time and verify that corrective actions have been effective at improving performance.**
3. **Develop performance attributes and evaluation criteria to assess performance of the composite adversary force (CAF). Incorporate these performance attributes and evaluation criteria into inspection procedures, assessment tools (including SDP and industry feedback), and guidance documents. Document assessments in inspection reports.**
4. **Ensure program guidance is adhered to such that a stable, predictable, and transparent regulatory path is consistently applied across the fleet of power reactor sites.**
5. **Evaluate additional methods of improving personnel safety during the conduct of FOF exercises to preclude inadvertent firing of live weapons.**
6. **Institute a formal, periodic self-assessment process (performance goals, evaluation criteria or standards, performance metrics, and periodic review) to ensure that the FOF program is accomplishing its objectives.**

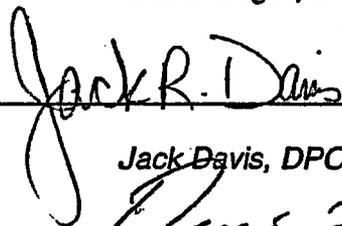
The Panel would like to acknowledge the professionalism, dedication, and tenacity of the submitter in pursuing his concerns. By raising them to this level, he has ensured that they will receive a full vetting. In this way, he has contributed significantly to the NRC's role of ensuring public health, safety and security. The Panel recommends that the submitter be considered for specific performance recognition in honor of his contribution.

Differing Professional Opinion (DPO)
on
Force-on-Force Evaluation Criteria

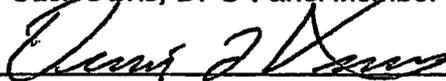
DPO Panel's Final Report



Bruce Boger, DPO Panel Chair



Jack Davis, DPO Panel Member



Dennis Vernon, DPO Panel Member



Rani Franovich, DPO Panel Assistant

Table of Contents

I.	Introduction	4
II.	Evolution of the Current Process	5
III.	The Current Approach	7
	1. The NRC's Use of a Win/Lose Paradigm	7
	2. Inconsistent Information and Staff Understanding	8
	a. Objective of the FOF Program	9
	b. Inconsistently Expressed Role of NRC in FOF Exercises	10
	c. Inconsistency with Reactor Oversight Process (ROP)	11
	3. Unclear Information and Staff Understanding	15
	a. Interface with Physical Protection Baseline Inspection Program	15
	b. Training and Certification of Evaluators	15
	4. Summary of Panel's Assessment of the Current FOF Approach	15
IV.	Planned Changes/Improvements to the Current Approach	16
V.	Comparison to DOE Approach	16
	1. Evolution of the DOE Rating System	17
	2. DOE Inspection Approach and Scope	18
	3. Current DOE FOF Program	19
	4. FOF Evaluators and Evaluations (resources)	20
	5. Evaluator Training	20
	6. Multiple Integrated Laser Engagement System (MILES)	21
	7. Personnel Safety	21
	8. Summary of Comparison with DOE's Approach	22
VI.	Timing of DPO Submittal	23

VII. Conclusions 24

VIII. Recommendations 24

IX. Closing Comments 26

Appendix 27

Disclaimer: Information contained in portions of this report reflects the DPO ad hoc review panel's understanding of the Department of Energy's (DOE) force-on-force program. In the interest of issuing the report in a timely manner, the DOE has not been afforded an opportunity to review and comment.

Report Details

I Introduction

An ad hoc review panel was established by memorandum dated January 24, 2005, to review a differing professional opinion (DPO) involving Force-on-Force (FOF) evaluation criteria. Bruce Boger served as the Chairman of the ad hoc review panel (hereafter the "Panel"). Jack Davis and Dennis Vernon, both from NSIR/DNS, served as panel members, and Rani Franovich of NRR/DIPM assisted the Panel.

The Panel reviewed the written concerns submitted in the DPO and summarized them as follows:

In short, the submitter contends that a win/lose paradigm is employed by the NRC in evaluating Force-on-Force (FOF) exercises at power reactor facilities. The submitter also contends that the win/lose approach is insufficient in that:

- (1) it does not adequately address relevant insights into licensee performance during the conduct of FOF exercises (i.e., is not performance-based);
- (2) results derived from these win/lose criteria are not reliable and do not provide defensible results and actions taken, or not taken, by the NRC and respective licensees due to these results may be detrimental to the public health and safety.
- (3) there are numerous artificialities, many of which are complex and significant, associated with these exercises and these artificialities routinely affect the outcome of exercises. Most of the consequences resulting from these artificialities are known and can be adequately factored into the evaluation process. However, there are consequences created by some artificialities that evaluators are either unaware of or do not fully understand and these consequences can be the determining factor in which side wins or loses an exercise
- (4) the affect of drill artificialities on the final outcome (success or failure to protect the target set) of the FOF exercises is amplified in the win/lose paradigm.

Regardless of the criteria used artificialities will always affect the outcome of these types of exercises. The number and type of artificialities, and the subsequent consequences relating to the bottom-line, remain the same regardless of the evaluation criteria used. However, a performance based criteria that evaluates the capabilities of protective force attributes such as Command and Control, Communications, Tactics (individual and team), Equipment, Protection Strategy, etc. along with other supporting protective systems is not reliant on the bottom-line results to obtain necessary and beneficial data. Using performance based criteria evaluators are able to determine the effectiveness of these individual attributes regardless of who wins or loses the exercise. Evaluating these essential attributes provide evaluators with information pertaining to the protective force/systems' effectiveness which in turn provides a basis for determining whether the licensee has the necessary capabilities to adequately protect essential equipment.

(5) it precludes adequate evaluations of "wins" to identify performance deficiencies and areas for improvement;

(6) it provides no incentive to licensees to function in an open, legitimate manner and may provide motivation to manipulate their performance in an effort to "win."

In the formal DPO, the submitter proposed a remedy for his concerns. The proposal is summarized as follows:

The submitter proposed that the NRC adopt a different evaluation approach similar to one employed by the Department of Energy. The proposed methodology, as the Panel understands it, is as follows:

Revise the FOF evaluation criteria. Employ performance-based criteria that focus on security force attributes (e.g., command and control, tactics, communications, etc.). An appropriate number of NRC evaluators with appropriate training should be used to assess protective force performance. Achievement of target set(s) provides insights into protective force performance, but the focus of performance assessment should be on the security force attributes and supporting protective systems. The focus of assessment should not be on win/lose. Win/lose is important to the assessment but the focus should be on why the licensee won or lost with emphasis on performance-based criteria.

In summary, the submitter contends that when using the Win/Lose criteria a licensee with a less than adequate protective system/force can "win" exercises while an licensee with adequate protective measures can "lose" due the existence of artificialities. This is especially significant when evaluators are either unaware of the introduction of one or more artificialities or the potential consequences of these artificialities. Moreover, the perceptions and actions taken based on these indefensible results can be counterproductive to the security posture at the evaluated facility and detrimental to credibility of the NSIR Force-on-Force program.

The Panel's summary of the concerns and proposal expressed in the formal DPO was approved by the submitter on April 11, 2005.

The objective of the Panel was to conduct a thorough review of the submitter's concerns and proposal and make a recommendation regarding the disposition of the issues presented therein. To this end, the Panel reviewed a number of NRC and DOE documents. The Panel also interviewed a number of NSIR/DNS staff and management, as well as representatives of DOE responsible for the DOE FOF program. This report documents the Panel's observations and conclusions based upon the information it reviewed.

II Evolution of the Current Process

The Panel thought it was important to understand the context within which the FOF Program has evolved over the last four decades. In 1967, the Atomic Energy Commission (AEC), interpreted its "common defense and security" legislative authority as mainly requiring basic protection against industrial sabotage at licensed commercial nuclear power plants (NPP). The AEC concluded that protection of the U.S. against hostile enemy acts was the responsibility of the nation's defense establishment and of the various agencies having internal security functions. However, there were numerous public challenges to this limited interpretation, most

notably during the Turkey Point NPP licensing hearings of 1967, that prompted the AEC to commission a study to investigate the appropriate scope of nuclear power plant industrial sabotage protection. It was determined that industrial sabotage should be viewed in terms of socioeconomically-motivated actions (such as acts resulting from civil demonstrations) and should not be defined as damage that could result from sabotage by subversive organizations or trained foreign agents. Such decisions were codified and promulgated in 10 CFR 50.13, "Enemies of the United States," which specifically precludes the requirement for NRC licensees to protect against sabotage from an enemy of the United States, whether a foreign government or other person.

Since 1967, there has been considerable change in AEC, and subsequently NRC, policy and licensing practice regarding security requirements for commercial nuclear power reactors primarily due to the change in the incidence and type of malevolent acts in society. Considerations of design threat levels as applied to reactor physical protection were included in the 1973 ANSI Standard N18.17, "Industrial Security for Nuclear Power Plants," and endorsed by Regulatory Guide 1.17 that same year. In early 1974, the security plans at all operating reactors were reviewed in accordance with the guidance of Regulatory Guide 1.17.

After the creation of the NRC in 1975, Congress directed the NRC to develop contingency plans for dealing with threats, thefts, and sabotage at NRC licensed facilities. By 1977, the NRC was directing its licensees to examine facility operations to determine vulnerabilities related to radiological sabotage. However, the NRC became concerned that implementation was inconsistent among licensees and in early 1982, the Regulatory Effectiveness Reviews (RER) Program was established. The intent of the RER program was to examine the effectiveness of the security requirements while evaluating the licensees' ability to implement the requirements in a manner that would ensure protection against the Design Basis Threat (DBT). Unfortunately, no implementing guidance existed at the time and this contributed to further inconsistencies in the findings from the NRC reviews. Thus, after the completion of several RER inspections the NRC employed external security experts to provide input to the review process that was consistent with the known threat environment and with known and expected adversary characteristics of the time.

As a result of the numerous deficiencies and inconsistencies in licensee security programs that were identified through the RER program, the NRC determined that it was appropriate to test licensee defensive strategies under simulated adversary attacks. In the late 1980's, the NRC established the Operational Safeguards Response Evaluation (OSRE) Program. To improve consistency in program application, the NRC developed target sets with thresholds based upon a Part 100 release. The NRC also established an explosives testing program to validate barrier time credits as used in the OSRE program. The OSRE program was sized for an 8-year frequency of NRC observed plant FOF testing and completed its first cycle in 2000.

In response to the events of September 11th the Commission temporarily suspended FOF evaluations. When the Commission approved staff plans to resume FOF exercises, it did so in an orderly phased approach. Phase 1 analyzed the impact of expanded adversary characteristics reflected in the new threat level via expanded tabletop drills. Phase 2 factored the lessons learned in Phase 1 into guidance governing pilot expanded FOF exercises. In addition to evaluating the process and scope of FOF exercises, the Phase 2 activities also evaluated the impact of the expanded adversary characteristics and compensatory measures.

Phase 2 was expected to be the final phase of the analysis of the FOF process before commencing a new, routine FOF exercise program. However, experience from Phase 2 convinced the NRC that further development was warranted to achieve more effective, efficient, and realistic FOF exercises. It was also important to incorporate the design basis threat (DBT) order into any routine FOF program. Thus, the NRC added a third Phase to complete transitional FOF exercises before a final routine program could be commenced.

The third Phase (Transitional Phase) continued to evaluate the purpose and scope of the future FOF program and the use of Multiple Integrated Laser Engagement System (MILES) equipment. It also explored the potential utility of software programs (e.g., Joint Conflict and Tactical Simulation [JCATS]) in preparing for and supplementing FOF exercises and piloted a significance determination methodology that considers mitigative actions and Part 100 criteria in assessing the potential consequences of adversary attacks. In addition, Phase 3 continued to expand the incorporation of Operations and Emergency Planning elements. Similar to Phase 1 and 2, the transitional phase was conducted under the condition that findings were not subject to the Interim Physical Protection Significance Determination Process (PPSDP) or enforcement as the program exceeded current regulatory requirements. Phase 3 was completed in October 2004 with the results and lessons learned from this phase documented in SECY-04-0174. Despite the fact that numerous components of the FOF program and necessary related components (e.g., the PPSDP) were still in draft, the staff did not identify any significant program issues to prevent transition into a full (routine) FOF evaluation program. Thus, the routine program commenced in November 2004.

Since the FOF testing program has developed in an incremental and evolutionary process over many years, the current program was not able to benefit from a comprehensive, structured, and measured foundation. As a result, improvements and associated results determined during the phased resumption have no baseline from which to substantiate the appropriateness of the program's essential elements (e.g., testing frequencies, minimum number of controllers/evaluators, win/lose paradigm, etc.) and made it challenging for the Panel to fully evaluate the merits of changes proposed by the DPO submitter without considerably more in-depth review that would result from a comprehensive program review.

III The Current Approach

The Panel reviewed documents and conducted interviews of DNS staff and management to gain a better understanding of the NRC's current FOF Program. Specifically, the Panel sought to confirm the submitter's assertion that the current FOF Program employs a win/lose paradigm. Based on its research and interviews, the Panel found that agency documents (inspection procedures, information on the website, and correspondence with internal and external stakeholders) at times provide inconsistent or unclear information about the FOF Program. The Panel also found that FOF Program inspectors and management do not share the same, consistent understanding of FOF Program objectives or inspection program guidance and framework. A detailed discussion of the Panel's findings is provided below.

III.1. The NRC's Use of a Win/Lose Paradigm

The Panel sought to confirm the submitter's assertion that the FOF Program applies win/lose criteria to evaluate protective force performance. The Panel reviewed inspection program

guidance documents and information on the NRC's website. The Panel also interviewed NSIR/DNS staff and management to glean their understanding of the current FOF evaluation approach and criteria used by the NRC.

Throughout NRC Inspection Procedure (IP) 71130.03, "Contingency Response - Force-on-Force Testing," the emphasis is clearly on the loss or successful protection of a target set. Specifically, Section 05.34, "Screening of Findings through SDP," focuses on target set losses and the impacts of radiological release as the sole factors considered in significance determination.

Similarly, Manual Chapter 0609, Appendix E, "Physical Protection Significance Determination Process for Power Reactors," focuses on target set loss or protection and radiological release in determining a finding's significance. The Panel noted that this focus on protecting (or losing) target sets applied to both the Interim PPSDP and the proposed FOF SDP.

In contrast, the NRC's fact sheet backgrounder on FOF exercises states the following:

The Nuclear Regulatory Commission (NRC) has carried out force-on-force (FOF) exercises regularly at commercial operating nuclear power plants since 1991 as part of its comprehensive security program. However, they are not pass/fail inspections.
<http://www.nrc.gov/reading-rm/doc-collections/fact-sheets/force-on-force.html>

The FOF SDP and the information provided on the NRC's public website appear to conflict because the former focuses on bottom-line outcomes while the latter implies that a more graded approach is applied to performance assessment.

During interviews, NSIR/DNS staff and management confirmed (without exception) that the NRC's FOF Program and evaluation paradigm revolves around the bottom line outcomes (protection of target sets) of the exercises. Specifically, the premise of the approach is that significant performance issues will manifest themselves in target set losses. The submitter contends that artificialities, both known and not recognized, may mask significant performance issues regardless of target set protection.

Based on its review of the program documents and interviews, the Panel concluded that the NRC's FOF Program does indeed apply win/lose criteria. The Panel also noted an apparent inconsistency between this conclusion and the statement that "they are not pass/fail inspections." Similar inconsistencies in other aspects of the program were observed and are discussed in the following Section.

III.2. Inconsistent Information and Staff Understanding

The Panel identified aspects of the FOF Program that appeared to be inconsistently documented and not mutually understood by NSIR/DNS staff and management. Specifically, inconsistencies were noted in the objective of the FOF Program; in the NRC's role in FOF exercises; and in the application of the philosophy, framework, and guidance documents of the Reactor Oversight Process (ROP). Examples of these inconsistencies are discussed in detail below.

III.2.a. Objective of the FOF Program

In evaluating the merits of the submitter's concerns, the Panel sought to determine if the objective of the NRC's FOF Program was to assess licensee performance or improve protective force capabilities. The Panel considered this distinction important because the submitter's proposal to focus on and evaluate performance attributes would appear to be more effective in improving performance than the win/lose paradigm. However, if the role of the FOF Program was to simply evaluate performance, then performance feedback at this level might not add sufficient value for the assumed additional cost in evaluator resources, training and certification.

NRC IP 71130.03 indicates that the overall objective is to assess performance. Specifically, the objectives of the inspection are, in part, to verify and assess the capability of licensees' physical protective systems and security organizations, their protective strategy, the implementation of their protective strategy, and conduct of FOF exercises.

In contrast, the NRC's website indicates that FOF exercises are used to evaluate and improve licensee performance. Specifically, the Fact Sheet Backgrounder on FOF exercises states the following:

[Force-on-force exercises] are the primary means to evaluate and improve the effectiveness of plant security programs to prevent radiological sabotage as required by NRC regulations (10 CFR Part 73).
<http://www.nrc.gov/reading-rm/doc-collections/fact-sheets/force-on-force.html>

The question of the FOF Program's objective has come up in correspondence to the Commission as well. In SECY-04-0083, "Final Report on the Pilot Expanded Force-on-Force Exercise Program with Lessons Learned and Recommendations for Future Activities," dated May 14, 2004, the staff identified a significant policy issue regarding the objective of the FOF program. In the Commission Paper, the staff states:

The ... policy issue regards the overall objective and purpose of the FOF program: training/familiarization, performance assessment, or a combination. Force-on-force exercises have been recognized by the Department of Defense (DoD) and the Department of Energy (DOE) and many others as an excellent method of providing hands-on training in both offensive and defensive tactical methods. These same organizations also regard FOF exercises as the best performance-based method of assessing tactical strategies, including perimeter detection and assessment systems, and unit capability. The staff believes that with adequate procedures and controls and consideration for the artificialities associated with simulated activities, both objectives can be met with the NRC FOF exercise program.

During interviews with NSIR/DNS staff and management, the Panel concluded that there was not a consistent understanding of the FOF Program objectives among the staff. The Panel detected a reluctance to provide feedback in the form of observations to licensees following the conduct of FOF exercises. IP 71130.03 acknowledges that observations are to be shared or "passed on to security personnel." However, this was not being consistently done in an effort to improve performance. The prevailing view expressed to the Panel by NSIR/DNS staff and

management was that sharing observations was not appropriate because it was perceived as being subjective and constituted "consulting" with the licensee. The Panel asked each person interviewed what they thought of the submitter's proposal to shift focus from drill outcomes to assessment of performance attributes and impact of artificialities on target set loss or protection. Without exception, each responded that this kind of feedback was considered an expression of personal opinion regarding performance and was not appropriate. The predominant sentiment was that the inspection teams must avoid consulting with the licensee. The Panel found these sentiments to be indicative of an extreme avoidance of inspection practices from OSREs of the past. Since licensees had complained about the subjective nature of OSRE inspections, the NRC adopted the win/lose concept of categorical performance assessment. While this approach may be more simple to implement, it focuses on outcomes to the exclusion (formally, at least) of potentially significant insights into artificialities and their impact on performance (success as well as failure of the protective force). It also appears to have created an unfortunate inability to distinguish (among DNS staff and management) between consulting with the industry (i.e., inappropriately proposing corrective actions or solutions to challenges) and providing valuable observations and constructive feedback (i.e., appropriately seeking to improve performance). The Panel concluded that improving performance was not considered an objective of the NRC's FOF Program by the majority of the NSIR/DNS staff and managers.

However, one senior manager in NSIR/DNS indicated to the Panel that the objective of the FOF exercises was to establish a baseline capability across the fleet of power reactor sites to defend against, or thwart, the revised design basis threat (DBT). This senior manager acknowledged that, for some power reactor sites, this would require a higher standard of performance than that which previously would have been considered acceptable for the former DBT. This indicated to the Panel a desire to improve performance across the fleet to some acceptable baseline level.

The Panel concluded that the objective of the NRC's FOF program, in practical terms, is to both assess and improve protective force performance. However, this objective was not consistently articulated in program documents or the public website; nor was it consistently understood and implemented by DNS/NSIR staff and management, who appeared to be very concerned with "consulting" with licensees.

III.2.b. Inconsistently Expressed Role of NRC in FOF Exercises

Another aspect of the NRC's FOF Program that generated some confusion for the Panel involves the role of the NRC in the conduct of FOF exercises. IP 71130.03 states, in Section 05.06, "Entrance Briefing," that "the [inspection] team, including NRC contractors, will not play an active role in any exercise." Section 05.32, "Licensee Exercise Critique," of the IP also directs inspectors to "evaluate the licensee's utilization of exercises both as a training tool and as a means of self-auditing the protective strategy."

Conversely, the NRC's public website implies that the FOF exercises are owned and controlled by the NRC. Specifically, the Factsheet on FOF Security Exercises states:

The NRC plans the FOF exercises, runs the exercises, and evaluates the exercises.

NRC evaluates the plant operator's performance during FOF exercises; not the plant operator.

<http://www.nrc.gov/reading-rm/doc-collections/fact-sheets/force-on-force-fs.html>

The Panel also noted NRC documents that cite the efforts to work with industry to develop a composite adverse any force (CAF) that is trained to standards issued by the Commission and that will be evaluated at each exercise using vigorous NRC performance standards issued in April 2004. The Panel did not see evidence of this evaluation in the FOF inspection procedure or FOF inspection reports. The Panel was not sure how the NRC accomplishes this not in FOF exercises.

The Panel sought clarification from NSIR/DNS staff and management during interviews. NSIR/DNS staff and management indicated that licensees are expected to perform FOF drills throughout the year, with or without NRC participation, and critically assess their own performance with a desire for continuous self-improvement. Nevertheless, when the NRC was onsite, it was the NRC's exercise. Based on the contradictions between the inspection procedure and the information provided during interviews, the Panel determined that there was confusion surrounding the NRC's role in FOF exercises.

III.2.c. Inconsistency with Reactor Oversight Process (ROP)

According to Inspection Manual Chapter (IMC) 0305, "Operating Reactor Assessment Program," the ROP "integrates the NRC's inspection, assessment, and enforcement programs. The Operating Reactor Assessment Program evaluates the overall safety performance of operating commercial nuclear reactors and communicates those results to licensee management, members of the public, and other government agencies. The assessment program collects information from inspections and performance indicators in order to enable the agency to arrive at objective conclusions about the licensee's safety performance. Based on this assessment information, the NRC determines the appropriate level of agency response, including supplemental inspection and pertinent regulatory actions ranging from management meetings up to and including orders for plant shutdown. The assessment information and agency response are then communicated to the public. Follow-up agency actions, as applicable, are conducted to ensure that the corrective actions designed to address performance weaknesses were effective."

In SECY 04-0198, "Redeveloping the Assessment Process for the Physical Protection Cornerstone of the Reactor oversight process," transmitted to the Commission on October 22, 2004, the staff informed the Commission of its intentions and direction in implementing Option 5 of SECY-04-0020, "Treatment of Physical Protection Under the Reactor Oversight Process," dated February 5, 2004. The staff indicated the following:

The staff considered what elements an assessment process must have to be considered as part of, or parallel to, the ROP. The staff determined that such a process should have the same basic objectives as the ROP, contain the same basic components as the ROP, and use objective measures to determine agency response based on licensee performance. The ROP is a regulatory framework that includes licensee performance indicator data, NRC inspections, determinations of significance of inspection findings, and assessment based on

the combination of PIs and inspection data. The goals of the process are for it to be objective, risk-informed, understandable, and predictable.

SECY-04-0198 further stated:

The basic elements of the ROP that the physical protection assessment process would need to maintain are:

- PIs and inspection findings that feed into an assessment program
- a defined, repeatable, objective process to determine the significance of inspection findings

In SECY-05-0082, "Revised Assessment Process for the Security Cornerstone of the Reactor Oversight Process," transmitted to the Commission on May 12, 2005, the staff indicated the following:

The staff followed the process described in SECY-04-0198 in developing a process to assess security performance of power reactor licensees.... Objectives of developing this process included providing valid indications of licensee performance, integrating with the ROP, and minimizing as much as possible impact on the regions. That is, the new process should use as much of the existing programs as possible to prevent duplication or performing the same function in different ways. The process the staff developed is based on the ROP and follows ROP principles, but separates the security-related assessment from the assessments of the other six cornerstones. The new process uses the same assessment inputs (i.e., inspection findings and performance indicators) as the ROP and combines the inputs in a separate action matrix to determine the appropriate agency response.

The staff reviewed recent FOF inspection reports to see how the assessments were developed and documented. One particular inspection report documented significant performance issues. However, only one green non-cited violation was identified for failure to adequately develop target sets in accordance with the security plan. The inspection report indicates that two target sets were protected and one target set was destroyed. The Panel understands through interviews that protracted discussions among the team members, team leader, and DNS management ensued, and a conclusion was reached to characterize the exercise as indeterminate. However, senior NSIR officials promptly visited the site to assess the situation. Although significant discussion of the artificialities and their impact on the exercise outcome was provided in the inspection report, neither the destroyed target set nor the performance deficiencies were developed as findings and evaluated in accordance with the FOF SDP.

It appeared to the Panel that the win/lose criterion was not applied in this case; rather, expert judgement and management discretion were applied, resulting in an outcome (including a site visit by senior NSIR/DNS managers and plans for a follow-up inspection) that was outside the formal process (i.e., not in accordance with IMC 0612 and the FOF SDP). The extent to which the formal assessment process relies on expert judgement is described in the FOF inspection guidance. Section 05.21, "Inspection Team Meeting," of IP 71130.03 states:

An adverse conclusion could be either a finding or an observation, which may only need to be passed on to security personnel. The team should reach a consensus on any potential finding and its relative importance.

However, IP 71130.03 does not offer guidance on what constitutes a finding; the distinction between a finding and an observation; or how a finding is developed and dispositioned. Rather, IMC 0612 is the governing ROP guidance document for issue screening. According to IMC 0612, a finding is defined as:

An issue of concern that is related to a licensee performance deficiency. Findings may or may not be related to regulatory requirements and, therefore, may or may not be related to a violation.

Although the SDP for FOF Exercises focuses on target set loss or protection and radiological release in determining a finding's significance, it refers the NRC inspector to IMC 0612, Appendix B, "Issue Screening," for detailed analysis of an issue. The SDP appears to categorize these "issues" as FOF Process Deficiencies and characterize them as green findings. The SDP defines a finding as:

An issue with some significance that has been placed in context and determined either to be of sufficient significance to warrant more detailed analysis using the SDP or to have extenuating circumstances warranting its documentation in an inspection report. To be a finding, it must pass through the threshold screening process described in MC 0612. Findings may or may not be related to requirements including requirements imposed by order.

As such, the Panel would have expected the inspection report to have included multiple findings, not just the one green NCV for exercise process issues involving selection of a target set. However, performance deficiencies during this FOF exercise were not identified, developed, characterized using a significance determination process, and documented in the associated inspection report as findings. Even the one lost target set identified in the recent FOF exercise did not result in an inspection finding. Similarly, the level of regulatory response was not commensurate with the documented findings. In the Panel's view, this was not in conformance with the ROP.

During interviews, DNS staff and management confirmed their reliance on expert judgement to assess protective force performance and identify findings. The Panel determined that, in practice, findings were specifically tied to lost target sets. In one interview, the Panel was told that findings must have a regulatory basis and that any compliance issue identified during a FOF exercise would not be characterized and developed as a finding unless one or more target sets were destroyed. A senior DNS manager confirmed that this was the practice. As such, it became apparent to the Panel that performance deficiencies would not be developed as FOF inspection findings, documented in inspection reports, communicated to licensees for corrective action, and revisited during future inspections to ensure they were adequately corrected, unless target sets were lost. In the Panel's view, this also was not in conformance with the ROP.

During interviews, DNS staff and management alike indicated that performance issues are not considered findings unless they are associated with multiple target set losses and the inherent

failure to protect against the Design Basis Threat (DBT). Even violations of NRC requirements might not be considered findings unless they involved target set losses. When questioned about this apparent deviation from the philosophy of the ROP, a senior DNS manager indicated that he was not so sure that the FOF should aspire to the same philosophy with respect to identifying and developing findings. In his view, the fundamental regulatory requirement that governs FOF exercises is the ability to thwart the DBT in accordance with the associated post-9/11 security order. He added that the ability to meet this regulatory requirement rested on the ability to protect target sets. If a licensee's protective force was successful in this regard, then the regulatory requirement would be met and no findings would be identified. The Panel showed the DNS senior manager the FOF SDP definition of "finding," which states that findings may or may not be related to requirements including requirements imposed by order. When questioned how his view of a finding is reconciled with the ROP philosophy, inspection program guidance documents (e.g., MC 0612), and the FOF SDP, he reiterated that the FOF program did not necessarily apply the same philosophy.

The premise of the current FOF program is that significant performance issues will result in target set losses. The Panel considered this premise to be flawed because performance issues may not necessarily translate into drill outcomes if they are neutralized or masked by artificialities. The Panel also believes that one particular exercise was characterized as "indeterminate" for this very reason. An artificiality (adversary theoretically neutralized himself before destroying a target set) was claimed by the licensee, and the corresponding assessment was rendered indeterminate. Performance issues were offset by fortuitous circumstances, and the inspection report did not produce findings that accurately represented licensee performance. Still further, the agency response was not commensurate with the findings documented in the inspection report. As such, the FOF program does not ensure that assessment is valid, repeatable, and consistent in accordance with the ROP and Commission Papers. Beyond the win/lose criterion for target set protection, expert judgement and management discretion are applied to assess the impact of artificialities on protective force performance. The process by which expert judgement and management discretion are applied is not formally established and proceduralized such that consistent outcomes are ensured. The Panel considers this to be contrary to the tenets of the Reactor Oversight Process assessment and response decisions were made without the benefit of "a defined, repeatable, objective process to determine the significance of inspection findings" that are "[fed] into an assessment program" in accordance with the ROP Framework Discussion in SECY-04-0198.

SECY-05-0082 states that an objective of the new security oversight process is to provide valid indications of licensee performance. In reviewing the Commission Paper, it appeared to the Panel that FOF exercise outcomes (target set defense or loss) are a dominant input to the Security Cornerstone Action Matrix. In light of the importance of this input in determining the level of agency response, the Panel concluded that the current FOF program (as defined in program guidance documents) is potentially susceptible to unreliable or incomplete conclusions regarding licensee (protective force) performance because it focuses on bottom line outcomes. As such, the NRC's ability to maintain an appropriate level of oversight might be vulnerable to invalid assessments of protective force performance during FOF exercises. Specifically, a protective force that protects target sets fortuitously rather than through a strong defensive strategy and robust skills and abilities might not receive the appropriate level of NRC oversight to effect improvement. Likewise, a protective force with a strong defensive strategy and robust

skills and abilities might lose a target set because of the impact of artificialities on drill outcomes and could receive a heightened level of assessment that is unwarranted and inefficient.

The Panel concluded that IP 71130.03 addresses the licensee's responsibility to limit artificialities before initiation of a FOF exercise, but it does not provide guidance on how to consider artificialities that are identified during an exercise. The focus of the current program is on the bottom line: win/lose. As such, inspection findings are not identified to address performance issues unless target sets are lost. Expert judgement and management discretion are applied to determine an appropriate agency response when poor performance is exhibited. And the goals of having a process that is "objective, risk-informed, understandable, and predictable" cannot be consistently ensured. This is not compatible with the ROP philosophy, framework, or guidance documents.

III.3. Unclear Information and Staff Understanding

The Panel came across a number of aspects of the FOF Program that were not clearly understood or well-documented. Specifically, the Panel was not sure how the FOF inspection interfaced with the Physical Protection Baseline Inspection program. The Panel also was not sure how FOF evaluators who are interim certified would receive final certification.

III.3.a. Interface with Physical Protection Baseline Inspection Program

In its review of the two most recent FOF inspection reports, the Panel noted that several security program/process findings identified during FOF exercises were characterized as unresolved items. According to DNS staff and management interviewed, these items would be followed up and resolved by regional security inspectors during subsequent baseline inspections. This interface, however, was not apparent in the inspection guidance documents reviewed by the Panel. As such, the process for transferring inspection findings from the FOF exercise inspection to a baseline inspection was not clear.

III.3.b. Training and Certification of Evaluators

The Panel received information related to the training and qualification requirements for FOF team leaders, safeguards specialists, and reactor engineers. The Panel did not question and development of these requirements, but noted that they are considered to be interim. It was not clear what process would be used to receive final certification. The Panel understands that future efforts are envisioned in this area, but at this time it does not appear that a structured process similar to the one employed in the development of reactor inspector qualifications in IMC-1245 was used to develop the requirements. The panel notes that specialized skills are necessary to evaluate protective force performance attributes that are beyond a win/lose assessment. Some of these skills are perishable and need refresher training.

III.4. Summary of Panel's Assessment of the Current FOF Approach

The Panel investigated whether any planned changes or improvements were being considered beyond the phased resumption that may have a bearing on the submitter's concerns. The Panel was informed by NSIR/DNS management that some of the submitter's concerns were being considered for future improvements to the FOF program. During the phased resumption

of the FOF program, NRC and industry implemented a number of enhancements to reduce artificialities and to make the exercises more realistic including the use of MILES gear and improving mock adversary qualifications and preparation by including enhanced access to insider knowledge. However, aside from various changes made during the phased resumption NSIR/DNS management could not produce any documentation specifically addressing the issues as raised by the submitter. In fact, no documentation of the comprehensive vision or plan for orderly and predictable improvements to the FOF program or for the planned periodic effectiveness assessment of this program could be identified.

Discussions with various NSIR managers, FOF Team Leaders, and FOF staff have revealed that there are significant uncertainties and inconsistencies in understanding concerning the current state of the program, the logical and prioritized assessment of effective improvements to the program, and the schedule, timing and resources associated with the implementation of improvements to the program.

IV Planned Changes/Improvements to the Current Approach

As part of the phased approach, a number of enhancements were identified and implemented, and a number of further planned improvements to the FOF program are still envisioned by NSIR/DNS management.

For instance, as part of the pilot program, NRC and industry have implemented a number of enhancements to make the exercises more realistic including use of Multiple Integrated Laser Engagement System (MILES) gear to enhance the realism of the weapons and combat and improvements in mock adversary qualifications and preparation by including enhanced access to insiders. In calendar year 2000, NSIR began to investigate the methods to bring technology-based solutions to the time, labor, and travel intensive portions of FOF planning and execution.

Although these improvements are discussed in various documents, the Panel could not locate any comprehensive vision and plan for orderly and predictable improvements to the FOF program or for the periodic effectiveness assessment of this program. Discussions with various responsible NSIR managers, FOF Team Leaders, and FOF staff have revealed that there are significant uncertainties concerning the current state of the program; the logical and prioritized assessment of effective improvements to the program; and the schedule, timing and resources associated with the implementation of phased improvements to the program.

V Comparison to DOE Approach

The Submitter believes that the impact of drill artificialities on the final outcome (success or failure to protect the target set) of the FOF exercises is amplified in the win/lose paradigm. The Submitter proposed that the NRC adopt an approach similar to one employed by the U.S. Department of Energy (DOE). To become more knowledgeable about the DOE's FOF Program, the Panel met with the DOE Office of Independent Oversight and Performance Assurance (OA). This enabled the Panel to gain insights into DOE's inspection methodology and, moreover, their FOF evaluation rating system.

V.1. Evolution of the DOE Rating System

In the early 1980's, DOE employed the win/lose paradigm as the driving factor for assigning a rating (Satisfactory, Marginal and Unsatisfactory) to describe the adequacy of the protection afforded to the "Protection of Special Nuclear Material (SNM)" as the major topical area of inspection. Other major topical areas included Program Protection Management and Protection of Information. Among the program elements that were evaluated under the major topical area of the Protection of SNM included physical security systems and, protective forces. Although there were even lower tier areas (e.g. training; performance of routine duties) under the subject area of protective force, the outcome of the FOF received the greatest visibility and served as the driving force for the rating assigned to the subject area of protective forces which subsequently, impacted the overall rating for the major topical area. As a result, a "win" on the FOF was deemed necessary to obtain a passing inspection grade. When a site did not "win," the DOE operations office manager (similar to a Regional Administrator) would be summoned to appear before a Congressional oversight subcommittee to explain why the site "failed" and what corrective actions were planned to ensure there would not be a recurrence. These hearings were closed to the public, and records of these meetings are classified. Invariably, DOE officials indicated that FOFs were not reliable indicators of the entire site's safeguards and security posture. The low reliability of the exercises was attributed to drill artificialities, which had a significant impact on the outcome of an exercise. Moreover, DOE field officials indicated that the effectiveness of the various protective elements or measures that were, in fact, a part of the site's protective strategy, but the effectiveness of those components were not duly considered and/or evaluated. As such, critical judgements on the adequacy of protection afforded to strategic national assets was based upon an imperfect performance evaluation tool under the "win/lose" paradigm that resulted in the diversion of funding and personnel resources.

An adversary relationship developed between the DOE Headquarters (that was conducting the inspections) and the DOE field offices because the site contractors were under pressure to "win the FOF at any cost" in order to avoid a Congressional oversight hearing. As a result, additional artificialities (e.g., gaming the exercise) interfered with the FOF evaluation process. Field office managers petitioned for a more balanced and accurate approach to characterizing the site's security posture, and the DOE's "win/lose" construct evolved to the current methodology. The "Office of Independent Oversight and Performance Assurance, Appraisal Process Protocols" dated January 2002, describes the current approach as follows:

If there are negative deficiencies, weaknesses, deficiencies, or standards that are not fully met, these must be considered individually and collectively and then balanced against any strengths or mitigating factors to determine the overall impact on the program's effectiveness. Factors that should be considered during analysis include:

- Whether the deficiency is isolated or systemic
- Whether program managers and other line managers knew of the deficiency, and if so, what actions were taken
 - The importance or significance of the standard affected by the deficiency

- Mitigating factors, such as the effectiveness of other program or program elements that may compensate for the deficiency
- The deficiency's actual or potential effect on mission performance or accomplishment
- The magnitude and significance of the actual or potential vulnerability to DOE interests resulting from the deficiency.

Based upon the analysis, a rating is assigned by the topical area team members and those ratings are: Effective Performance (Green); Needs Improvement (Yellow) or Significant Weakness (Red). These ratings are described in detail in the aforementioned procedure. Each topic team submits their draft report to a Quality Review Board to ensure it is readable, logical (facts support conclusions and conclusions support ratings), and contains adequate balance. The draft report then continues through a process that includes verification of factual accuracy conducted by both the DOE field office and the site contractor(s) prior to being finalized and further disseminated.

By contrast, the NRC (in practice) currently relies on expert judgement and management discretion to draw conclusions about a protective force's performance as the analytical (i.e. significant determination process) and associated procedures have not been finalized. As has been stated earlier in this report, the NRC's assessment of the protective force's performance in one recent FOF exercise was ultimately determined by management decision and not by a process.

V.2. DOE Inspection Approach and Scope

During a routine OA inspection, the following topical areas are evaluated: Protection Program Management; Physical Security Systems; Material Control and Accountability; Protective Force; Personnel Security; Classified Matter Protection and Control; Unclassified Cyber Security Program; and, Classified Cyber Security Program. A minimum of two DOE inspectors are assigned to each topic.

Typically, each topical area is assigned a narrative rating - Effective Performance (Green); Needs Improvement (Yellow) or Significant Weakness (Red). As an example, under the topic of Protection Program Management, the sub-topical area of Management Direction and Guidance for Program Implementation is assessed by evaluating the budget, policy and staffing. Other sub-topical areas include an evaluation of the local DOE field office assessment program.

Sub-topical areas for the Protective Force topic are: Management; Training; Equipment; Facilities; Routine Duties; and, Emergency Duties. The outcome of a FOF is not a sub-topical area that is rated but, conversely, the results of FOF testing may be used as data points for the aforementioned sub-topical areas under the Protective Force topic. For example, if DOE evaluators determine that the protective force was not adequately trained in individual or team tactical movements (taking advantage of cover and concealment, radio discipline, and/or firearms proficiency, etc.), these types of deficiencies could provide indications of performance under the sub-topical areas of "Training" or "Management." These deficiencies also could

impact the performance rating in the topic area of "Protection Program Management" if training was not, in fact, provided as a result of budget, policy and/or staffing shortfalls, which might indicate that management did not appropriately prioritize the security needs to ensure that the primary security mission could be accomplished. Additionally, deficiencies identified in the FOF may also indicate less than adequate performance for one or more sub-topical areas under Protection Program Management topic. This would be considered in the overall assessment of the Protective Force's performance.

The NRC, in practice, relies on expert judgement and management discretion to draw conclusions about a protective force's performance. For example, one recent FOF exercise involving particularly poor performance resulted in one green finding involving the identification of target sets. But other performance issues were noted in the inspection report. The protective force was constantly in pursuit of the adversary force. The site's protective strategy did not, in fact, provide concentric circles of protection (defense in depth) and/or a response strategy that would ensure interdiction (at strategic choke points/avenues of approach) to the target set. The licensee's protective force argued that the loss of a target set was actually a win because an adversary neutralized himself. This "win" could not be attributed to good protective force performance, yet no finding was identified. And no findings were developed to address these performance issues. As such, the NRC's assessment of the protective force's performance was ultimately determined by management decision and not by a process. In contrast, the DOE system is designed with an independent quality review panel that reviews the work products of each team prior to going to senior management. Additionally, the DOE team spends one week on site for data collection and the following week for report writing which is also conducted on-site.

The Panel found that the DOE has a mature program that has been time tested; a documented formality and structured approach for inspection activities; an analytical process that involves the input of trained DOE inspectors; and review by a separate and independent peer quality assurance panel. The DOE also has established clear management expectations that, during the inspection, data are validated to ensure factual accuracy so that the conclusions are valid and based on the facts that were derived from field activities.

V.3. Current DOE FOF Program

Within the DOE, there are multiple organizational levels that conduct FOF exercises for different purposes. The site protective force contractors conduct exercises for individual and team tactical training; for in-service tactical response training; and to validate accuracy for adversary interdiction times based upon computerized vulnerability assessment data. Additionally, DOE requires their contractors (as a portion of their self-assessment program) to conduct four (4) equally-distributed FOF exercises each year. The technical basis for four (4) FOF's is to ensure that each protective force shift is tested. As a matter of routine, the local DOE field office during routine inspections conducts FOF exercises to verify the contractor's response capability as documented in the SSSP. The DOE line management organization may also conduct a FOF for quality assurance purposes to ensure that SSSP integrity is accurate. Lastly, the Office of Independent Oversight and Performance Assurance (OA) conducts FOF exercises as a part of their inspection data collection activities. The OA provides the Secretary of Energy with assurance of the adequacy of the site's FOF testing program and the site's ability to execute response plans.

It should be noted that only the Secretary of Energy can accept risk. Therefore, the technical basis for any site vulnerability should be documented in the SSSP so that the associated risk can be formally accepted by the Secretary. It should also be noted that OA conducts four (4) FOFs during their inspection (2 per night); NRC inspections conduct 3 FOF's (1 per night). DOE submits that since each exercise is different, results provide a shaky statistical foundation to support a particular generalization but a high degree of confidence to evaluate ad hoc planning in a tactical environment.

V.4. FOF Evaluators and Evaluations (resources)

FOF's are conducted under the auspices of the Protective Force Topic Team. However, DOE inspectors from other topical areas have been trained to serve as FOF evaluators and are deployed to serve in that capacity as needed. Therefore, at least 17 to 20 DOE inspectors are deployed as evaluators for a typical FOF. Their assignments may be either to a particular protective force armed responder (fixed or roving patrol posts) or to monitor activities in a particular sector or zone. After the exercise is completed, each DOE evaluator prepares a detailed "Office of Safeguards and Security Evaluations Evaluator's Worksheet" that documents their evaluation of the contractor's performance in such areas as planning; communications; command and control; individual tactical skills; team tactical skills; application of deadly force; and response plan execution. After the DOE evaluators attend the contractors comprehensive debriefing of the exercise, the DOE compiles the results and performs a trending analysis.

The Panel determined that the NRC had sponsored a study by Batelle, Columbus Division, on tactical exercise planning in the late 1980s. The results of the study were published in NUREG/CR 5081, "Tactical Exercise Planning Handbook," in April 1989. Section 6.1.1.2, "Evaluators," states that "exercise evaluators should be provided at the Central Alarm Stations (CAS) and Secondary Alarm Station (SAS), and with all major elements of the responding security forces. Additionally, at least one evaluator should be in a tower or other elevated position from which a general view of the exercise can be gained. Depending on the scenario, it may also be desirable to have an evaluator with the adversary force. Generally, six evaluators are sufficient for an exercise of this type." The Panel has learned that NRC FOF exercises involve the use of four NRC security evaluators, including the team leader.

V.5. Evaluator Training

DOE has a National Training Center (NTC) located in Albuquerque, New Mexico, that is used to develop and conduct all safeguards and security related training. One of the courses offered by the Center is designed for FOF evaluators. In comparison, NRC FOF personnel were not trained at the NTC; rather, they were trained by DOE personnel that provide secure ground transportation for special nuclear material and conduct FOF exercises to test their strategy to protect these safe, secure transport vehicles. As a part of their certification program, NRC FOF evaluators attended two external training courses at the Federal Law Enforcement Training Center (FLETC) and the DOE Office of Secure Transportation (OST). The training received from FLETC instructors was on the capabilities of the various firearms that the NRC staff may encounter during the conduct of a FOF evaluation. The weapons selected (Colt M - 4 carbine [5.56mm] and the 9 mm Sig Sauer P-226 handgun) were anticipated to be the type that the NRC staff may encounter while conducting inspections at power reactor sites. In order to gain an appreciation of the firepower of these weapons, each NRC person was afforded the

opportunity to fire the weapons at the firing range. At the conclusion of this block of instruction, a member of the NRC staff who has a military special operations training background, took advantage of the fortuitous work environment and demonstrated some individual and team tactical maneuvers. The NRC training group was then provided hands-on experience on tactical formations and movements while the FLETC instructors observed the training and offered some suggestions for performance improvement. All of the NRC trainees received a FLETC course completion certificate that states they have "attended the Nuclear Regulatory Commission Weapons and Tactics Familiarization Curriculum." That is the extent of the training provided to FOF evaluators from outside sources. The Panel noted that no internal NRC courses on combat skills (tactics, command and control, communications, etc.) are available.

The Panel is concerned that the training provided to NRC FOF evaluator may not be sufficient to provide the knowledge and skills needed to effectively and consistently evaluate protective force performance.

V.6. Multiple Integrated Laser Engagement System (MILES)

The NRC (Category I fuel fabrication facilities) and the DOE FOF Programs originated at approximately the same time (mid-1980's). The programs were developed by the same contractor (Battelle Memorial Institute, Columbus Ohio) and utilized essentially the same processes, protocols and procedures (Tactical Training Reference Manual, NUREG/CR-5172, and Tactical Exercise Planning Handbook, NUREG/CR-5172, both published April 1989). At this time, MILES equipment was, in fact, used during these Category I performance assessment exercises. In 1991, the OSRE Program for power reactor sites was initiated but the performance exercises were not MILES-enhanced. In lieu of MILES, rubber guns were used and controllers were forced to make decisions about kills and misses based on subjective judgment during the simulated adversary force engagement. In 2003, MILES was first introduced to the NRC's reactor FOF program.

V.7. Personnel Safety

Each DOE FOF exercise is governed by a formal Health and Safety Plan and any scenario-specific rules of conduct. DOE has a Safety Officer that accompanies the inspection team. The DOE Safety Officer serves in the role as the DOE Safety Controller and coordinates with the site safety staff to ensure that all safety aspects in preparation for the exercise are addressed. One of his functional responsibilities is to ensure completion of a Risk Analysis Report which documents a detailed risk/hazard analysis, as well as mitigating controls (i.e. use of appropriate personal protection equipment) for the force-on-force exercise. Foremost in a DOE FOF (or in a U.S. Department of Defense [DoD] FOF exercise for that matter) is the prohibition of live-fire weapons of any type and live-ammunition of any type in the test area.

DOE has imposed strict safety requirements that govern force-on-force exercises as a result of the tragic and preventable accidental shooting of a 36 year old man. On December 20, 1994, a fatal shooting accident occurred at the Los Alamos National Laboratory during a Engagement Simulation System (ESS) Limited Scope Performance Test. In essence, live ammunition was introduced into the performance testing area and allowed to be intermingled with blank ammunition. As a result of this accident, all ESS exercises across the Department were

suspended until provisions were instituted to prevent recurrence of a similar accident. Subsequently, the Department's firearms experts developed a Safety Panel that published a "Guide for Use of Protective Force Engagement Simulation Systems" dated April 10, 1995.

The DOE Safety Panel concluded there are two key elements to the conduct of such exercises in a safe manner. One is that the firearms be permanently modified to an approved configuration and dedicated to ESS use only. The second key element is to ensure that trained exercise controllers, in sufficient numbers, be employed to supervise and control every aspect of such exercises. All personnel that are assigned exercise controller or evaluator duties are required to receive formal documented training for the safe conduct of exercises, and the lesson plans were developed by the National Training Center (NTC). One of the requirements for the Shadow Force Controller is to ensure that "all live weapons are maintained under supervision, and Shadow Force personnel do not come in contact with exercise personnel with ESS equipment" ("Guide for Use of Protective Force Engagement Simulation Systems").

NRC allows the intermingling of both live weapons and live ammunition in the performance test area. The application of DOE lessons learned from this accident has not been applied to NRC force-on-force exercises due to the additional licensee cost for employing additional controllers for the shadow force. Unlike DOE, which pre-positions the shadow force with controllers at strategic response locations prior to the commencement of the FOF, NRC has both the shadow force member and FOF exercise participant intermingle under a single controller until the start of the exercise. When the exercise begins, the controller stays with the exercise participant while the shadow force member continues performing routine duties adding to the exercise artificialities as a distraction, at least, and possibly being misidentified as an exercise participant. NRC has required licensees to impose administrative controls (e.g. ribbon around trigger guard) with the intent to remind the armed responder that he is in possession of a live weapon. All of these exercise controls have been mandated by NRC without formal documentation or due consideration of personnel risk in light of other government agency lessons learned resulting from accidental shootings during FOF exercises.

The Panel found that NUREG/CR-5081 actually provided recommendations to address personnel safety. Section 6.1.1.7, "Shadow Force," states that "In general, a maximum of one shift is needed to cover actual security requirements for any given FOF exercise. These personnel form the shadow force that is positioned in one or more holding areas that are off-limits to problem play." Appendix E, "Force-on-Force Safety Plan," states under Section 2.1.1: "No live-fire weapons of any type will be allowed in the exercise play area, unless the personnel having such weapons are under the direct observation and supervision of a controller." Section 2.1.6 of this same Appendix E of NUREG/CR-5081 states "No live ammunition of any type or caliber will be brought into the exercise play area, unless it is under the direct supervision of a controller or is secured so as to be inaccessible during the exercise, unless an actual security emergency arises during such exercise."

V.8. Summary of Comparison with DOE's Approach

The DOE does not apply the win/lose criteria to its FOF exercises and focuses on evaluating combat skills (individual and team tactics; command, control, and communication; and use of deadly force) of the protective force in a simulated attack by a design basis threat adversary. The DOE's exercises are intensive, relatively frequent, and they involve a detailed approach to

assessment of performance attributes. The DOE also uses FOF exercises to assess the effectiveness of the site's performance testing and the to evaluate the suitability of site security plans. The Panel concluded that the DOE/AO's FOF Program appears to be very mature and generates valid, reliable outcomes. As such, the DOE representatives considered their program capable of providing a high degree of confidence in the effectiveness of the defensive strategies and capabilities of the security forces protecting its facilities.

By contrast, the NRC's FOF program is relatively new, it replaces the former OSRE process, it may not utilize an adequate number of evaluators, and it is still evolving. As such, the Panel believes that NRC can benefit from dialogue with DOE on the evolution of their program away from purely a win/lose paradigm to an approach that methodically and deliberately evaluates the skills and abilities of the protective force (focusing on critical performance attributes) and considers the impact of artificialities on drill outcomes. Although the Panel does not necessarily believe that the NRC should abandon its current program in favor of DOE's program, it does believe that the NRC can apply lessons learned by DOE and find merit in the evaluation tools developed by DOE. Efforts to benchmark with DOE could enable the NRC to augment its current evaluation criteria (defense of target sets) and thereby have a more complete understanding of a protective force's strengths and weaknesses such that the performance issues can be more specifically addressed. Efforts to further improve personnel safety during FOF exercises might also result from productive benchmarking with DOE.

VI Timing of DPO Submittal

As noted in MD 10.159, "In the free and open discussion of agency issues, professional differences of opinion are common. Employees normally try, and are encouraged, to resolve their concerns through discussions with their co-workers and supervisors." The MD also states that "In some cases, informal discussions fail to completely cover the matter in question ..." and if "all attempts to resolve the technical, legal, policy issues informally have failed," an employee may submit a formal DPO.

In the instance at hand, the submitter began employment with the NRC in 2003 with over 15 years of experience in the security field. In particular, he was very familiar with the FOF process used at DOE facilities, since he served as a team leader for FOF exercises while at DOE. He was hired to perform a similar role at the NRC. He participated in the development and transition of the current NRC FOF exercise program. The submitter was aware that DOE had modified their FOF program to reduce the emphasis on win/lose and to enhance the emphasis on protective force performance attributes. As explained below, he informally sought to introduce this approach to the NRC FOF program.

The submitter advised the Panel that he emailed a paper delineating his concerns to several co-workers and his first and second level supervisors in May 2004. He was unable to provide the Panel with a copy of that email, but indicated that it was similar in substance to the information provided in his DPO. His first-level supervisor recalled the email. The submitter indicated that, subsequent to that time, he spoke openly about raising his concerns through the DPO process. In addition, he met with his first level supervisor and discussed his concerns. This supervisor related to the Panel that he felt the approach proposed by the submitter would inappropriately place the NRC in a consultant role with the licensee. The concerns were next discussed between the submitter and his first and second level supervisor in the July 2004 time

frame. The second level supervisor indicated to the Panel that he saw some merit in the submitter's suggested approach, but he also felt that the approach tended to look like a consultant's role rather than a regulator. It does not appear to the Panel that the issues the submitter raised were elevated to higher management at that point in time. This may have been because the second level supervisor felt he had remanded the issues back to the first level supervisor for resolution. However, the first level supervisor seemed to feel that it was the submitter's prerogative to submit a DPO if further recourse was desired. The submitter recalls being asked about his intentions to submit a DPO by his first and second level supervisors after their meeting, but no follow up discussions to further pursue his issues took place.

In the Panel's view, the submitter's concerns were not well-communicated up the NSIR management chain prior to the submittal of the DPO. Although opportunities existed, a thorough vetting of the submitter's concerns did not take place to ensure that they were fully considered and dispositioned by NSIR management. The Panel recognized that the timing of the concerns raised by the submitter was challenging because the organization was focused on the implementation schedule for the revised FOF exercise program (November 2004). Many activities were taking place during the spring and summer of 2004 to update procedures and provide training to security inspection staff. Nevertheless, opportunities existed to more fully address the submitter's concerns.

VII Conclusions

1. The objective of the FOF Program is not clearly or consistently defined.
2. The FOF Program is not well defined with attendant procedural guidance to ensure a stable, predictable, transparent regulatory path is consistently applied across the fleet of power reactor sites. As a result, the formalized program is not effective at addressing performance issues that do not result in clear target set losses. Rather, an informal approach that relies on expert judgement and senior management review is used to address cases of poor performance.
3. Training and certification of NRC evaluators is considered interim and is not based on any job-task analysis to ensure a baseline level of knowledge, skills and abilities is applied during inspections. This is especially significant because of the heavy reliance on expert judgement.
4. The submitter's concerns have merit, and his proposal to evaluate performance attributes against evaluation criteria and consider the impact of artificialities on drill outcomes would improve the quality and reliability of the NRC's FOF program for protective force performance assessment.

VIII Recommendations

With respect to the NRC's evaluation of FOF exercises, the Panel recommends consideration of the following:

1. Benchmark with other federal agencies (DOD, DOE) to gather insights, techniques and strategies for evaluating protective force performance during FOF exercises.

2. Develop protective force performance attributes and evaluation criteria that should be used by FOF evaluators.
3. Incorporate protective force performance attributes and evaluation criteria into inspection procedures, assessment tools (including those for the SDP and licensee feedback), and guidance documents. Document assessments in inspection reports.
4. Develop a structured process to provide licensees FOF exercise feedback on protective force performance against the attributes and evaluation criteria identified.
5. Perform a job-task analysis for NRC FOF evaluators. Develop a more structured and specialized training and certification program (including refresher training) for all FOF evaluators.
6. Employ a Communication Plan to reach out to internal and external stakeholders to develop a consistent understanding of the revised program.

With respect to the Panel's general observations of the security cornerstone and oversight process, the Panel recommends consideration of the following:

1. Revisit the FOF program's objectives. Clearly and consistently articulate those objectives in program documents, procedures, and communications with internal and external stakeholders.
2. Ensure that FOF program documents are complete and provide guidance to ensure that performance issues are identified, developed into findings, and documented in inspection reports. This will enable the staff to trend performance over time and verify that corrective actions have been effective at improving performance.
3. Develop performance attributes and evaluation criteria to assess performance of the composite adversary force (CAF). Incorporate these performance attributes and evaluation criteria into inspection procedures, assessment tools (including SDP and industry feedback), and guidance documents. Document assessments in inspection reports.
4. Ensure program guidance is adhered to such that a stable, predictable, and transparent regulatory path is consistently applied across the fleet of power reactor sites.
5. Evaluate additional methods of improving personnel safety during the conduct of FOF exercises to preclude inadvertent firing of live weapons.
6. Institute a formal, periodic self-assessment process (performance goals, evaluation criteria or standards, performance metrics, and periodic review) to ensure that the FOF program is accomplishing its objectives.

IX Closing Comments

The Panel would like to acknowledge the professionalism, dedication, and tenacity of the submitter in pursuing his concerns. By raising them to this level, he has ensured that they will receive a full vetting. In this way, he has contributed significantly to the NRC's role of ensuring public health, safety and security. The Panel recommends that the submitter be considered for specific performance recognition in honor of his contribution.

**Appendix
List of Documents Reviewed**

- NRC Inspection Procedure 71130.03, Contingency Response – Force-on-Force Testing, November 10, 2004.
- NRC Inspection Procedure 71130.05, Protective Strategy Evaluation. Dated February 19, 2004.
- NRC Inspection Manual Chapter 0609, Appendix E, Physical Protection Significance Determination Process (SDP) for Power Reactors (interim and draft).
- NRC Inspection Manual Chapter 0305, "Operating Reactor Assessment Program," dated December 21, 2004.
- NUREG/CR-5081, Tactical Exercise Planning Handbook (Battelle).
- "Context and Protocols for Performance Testing of Protective Forces," (DOE, Office of Environment, Safety and Health), February 1999.
- "Security Performance Evaluation Section Force-on-Force Program Review," Slides from Presentation by Ronald Albert, Section Chief, DNS/SPES, September 6, 2004.
- SECY-04-0020, "Treatment of Physical Protection Under the Reactor Oversight Process," dated February 5, 2004.
- SECY-04-0083, "Final Report on the Pilot Expanded Force-on-Force Exercise Program with Lessons Learned and Recommendations for Future Activities," dated May 14, 2004.
- SECY 04-0198, "Redeveloping the Assessment Process for the Physical Protection Cornerstone of the Reactor oversight process," dated October 22, 2004.
- SECY-05-0082, "Revised Assessment Process for the Security Cornerstone of the Reactor Oversight Process," dated May 12, 2005.

Telfair, W.D., D.A. Moul, J.W. Klingelhofer, W.R. Leonard, ATactical Exercise Planning Handbook, @ NUREG/CR-5172, U.S. Nuclear Regulatory Commission (prepared as BMI-2166, by Battelle Columbus Division), Washington, DC, April 1989.

"Context and Protocols for Performance Testing of Protective Forces," dated February 1999, published by the Office of Oversight, Environment, Safety and Health

"Health and Safety Plan for Safeguards and Security Evaluations Inspection Force-on-Force Exercises Using Engagement Simulation System (ESS)/Multiple Integrated Laser Engagement Systems (MILES) Activities at U.S. Department of Energy Sites," dated September 2, 2004