REVISED: 9/7/2010

August 13, 2010

UNITED STATES OF AMERICA
NUCLEAR REGULATORY COMMISSION

BEFORE THE ATOMIC SAFETY AND LICENSING BOARD

In the Matter of	)	
	)	
NORTHERN STATES POWER COMPANY	)	Docket Nos. 50-282-LR/ 50-306-LR
	)	
(Prairie Island Nuclear Generating Plant,	)	
Units 1 and 2)	)	

NRC STAFF REBUTTAL TESTIMONY
OF DR. VALERIE E. BARNES, JUNE CAI, MOLLY JEAN KEEFE,
AND AUDREY L. KLETT CONCERNING SAFETY CULTURE AND
NRC SAFETY CULTURE POLICY DEVELOPMENT AND IMPLEMENTATION

Q1. Please state your name, occupation, and by whom you are employed.

A1(a). My name is Dr. Valerie E. Barnes ("Barnes").[1] I am employed by the U.S. Nuclear Regulatory Commission ("NRC") as Senior Technical Advisor in Human Factors, Office of Nuclear Regulatory Research. A statement of my professional qualifications is attached to my July 30, 2010 pre-filed testimony.

A1(b). My name is June Cai ("Cai"). I am employed by the NRC as the Senior Safety Culture Program Manager in the Office of Enforcement. A statement of my professional qualifications is attached to my July 30, 2010 pre-filed testimony.

A1(c). My name is Molly Jean Keefe ("Keefe"). I am employed by the NRC as a Human Factors Specialist in the Health Physics and Human Performance Branch of the Division of Inspection and Regional Support in the Office of Nuclear Reactor Regulation ("NRR"). A statement of my professional qualifications is attached to my July 30, 2010 pre-filed testimony.

A1(d). My name is Audrey L. Klett ("Klett"). I am employed by the NRC as a Reactor

[1] In this testimony, the sponsors of each numbered response are identified by their last name; no such designation is provided for paragraphs that are sponsored by all witnesses.

Operations Engineer in the Performance Assessment Branch of the Division of Inspection and Regional Support in NRR. A statement of my professional qualifications is attached to my July 30, 2010 pre-filed testimony.

Q2. What is the purpose of your testimony?

A2. (All) The purpose of this testimony is to respond to testimony and exhibits submitted by the PIIC and the Applicant (PINGP) on July 30, 2010.

Q3. What knowledge, training, and/or experience are necessary to offer an expert opinion on safety culture?

A3. (Barnes) In order to render a reliable expert opinion on the safety culture of a complex organization, a person would need to obtain a master's degree or a PhD in at least one of the following areas of the social sciences that study human behaviour in group settings: sociology; anthropology; or certain sub-disciplines within psychology, like social, organizational, and applied psychology. In addition to coursework, an expert needs to have some field experience under the supervision of a properly qualified professional in the aforementioned fields. A person with a PhD degree would normally have received this supervised field work during the course of his or her study. Persons with master's degrees would need at least one year of supervised field work.

Q4. Why is field work necessary for being able to render a reliable expert opinion regarding the safety culture of organizations?

A4. (Barnes) Although graduate-level education in one of these disciplines ensures that the individual is knowledgeable of the theory and research literature applicable to safety culture and related topics, appropriate research methods for studying and analyzing safety culture, the strengths and limitations of the different research and evaluation methods, and the professional ethics of the discipline, a key aspect of providing reliable opinions is to understand any likely biases that the researcher may introduce into a study or assessment project and appropriate methods for testing and overcoming them.

Q5. Is specific knowledge, training, or experience necessary to interpret the results of a safety culture assessment?

A5. (Barnes) Yes. The types of education, training, and experience I discussed above are necessary to interpret the results of a safety culture assessment. Accurate interpretation of results depends, for example, on the ability to evaluate the likely validity and reliability of the methods used, the weight to place on different sources and types of information gathered in the assessment to appropriately integrate information obtained from disparate sources, and knowing how to arrive at well-founded, defensible conclusions about the likely generalizability of the results.

Q6. In A11 of his testimony, Mr. Grimes describes his experience in nuclear reactor safety management. Is safety management the same as safety culture?

A6(a). (Barnes) "Safety management" and "safety culture" are not the same thing. While there is some disagreement regarding the definition of "safety management," INSAG-13 (NRC Staff Exhibit NRC000054) provides a widely accepted definition of safety management. INSAG-13 defines safety management as a necessary, but not sufficient, element of a positive safety culture. INSAG-13 (NRC Staff Exhibit NRC000054) states:

> The structural aspect of safety culture comprises the organization's arrangements for safety, which is commonly described as the safety management system of the organization. 'Management' is used to mean the administration of the organization. … The safety management system comprises those arrangements made by the organization for the management of safety in order to promote a strong safety culture and achieve good safety performance.

Thus, the concept of nuclear power plant safety management focuses on the programs and processes the organization implements to ensure safe operations, which are important components of safety culture, but are only one source of information about an organization's safety culture (see A6(b) Cai below).

The large majority of the NRC's regulations for and its inspections of nuclear power plants focus on safety management. NRC regulations in Title 10 of the Code of Federal

Regulations impose requirements on licensees, such as 10 CFR Part 26, “Fitness for Duty Programs,” or 10 CFR 50.65, “Requirements for monitoring the effectiveness of maintenance at nuclear power plants,” which licensees then implement through “Fitness-for-Duty” or “Maintenance Rule” programs and accompanying procedures. Inspectors evaluate the effectiveness of these programs, the extent to which licensees follow their own procedures, and the licensee’s root cause analyses of problems that may arise, which may provide additional insights into the functioning and effectiveness of these programs, i.e., the licensee’s “safety management systems.” Although monitoring of licensees’ programs and procedures is important for assuring safe operations, the Commission has recognized that public health and safety would be better assured by more fully addressing safety culture in its regulatory oversight. Specifically in response to lessons learned from the Davis-Besse event, the Commission directed the Staff to enhance the Reactor Oversight Process (ROP) to more fully address safety culture in SRM-SECY 04-0111, “Recommended Staff Actions Regarding Agency Guidance in the Areas of Safety Conscious Work Environment and Safety Culture (August 30, 2004) (NRC Staff Exhibit NRC000044).

A6(b). (Cai) See A6(a) Barnes above regarding the definition of safety management. In addition, enclosure 3 of NRC SECY 06-0122, “Safety Culture Initiative Activities to Enhance the Reactor Oversight Process and Outcome of the Initiatives,” (NRC Staff Exhibit NRC000055) provides a thorough discussion of the specific differences between safety culture and safety management, and what concepts are uniquely covered by safety culture:

> Dr. Edgar H. Schein, Sloan Fellows Professor of Management Emeritus, Massachusetts Institute of Technology (MIT), developed a model of culture which IAEA and others adopted for their safety culture model. Figure 2 [of NRC Staff Exhibit NRC000055 at Enclosure 3] is representative of Dr. Schein’s model which has three levels of organizational culture:
>
> • Level 1: Artefacts (visible): Tangible products, behavior, organization structure and production processes
> • Level 2: Espoused values (not visible, but can be elicited): Values and norms, strategies, objectives, philosophy

> • Level 3: Basic assumptions: Unconscious and self-evident beliefs and assumptions
>
> The concept of safety culture encompasses all three levels of Schein's model, including basic assumptions. Safety management elements primarily fall into Dr. Schein's Level 1 – the regulations and procedures (the *how* safety is achieved), Level 2 - policy statements, and the outcomes of safety management that can be observed in behaviors. Although an organization's basic assumptions are less amenable to direct regulatory oversight than visible artefacts and espoused values, problems at Level 3 can serve as the "root causes" that may lead to repetitive and far-reaching safety performance problems. The regulator can gain insights into the third level for plants exhibiting declining performance, through more intrusive oversight in supplemental reactive inspection programs. For example, through evaluating the licensee's safety culture assessment and through NRC's independent assessment of the licensee's safety culture in the conduct of IP 95003.
>
> . . . .
>
> The NRC's proposed components of safety culture contain a mix of Level 1: performance outcomes; Level 2: policies, programs and processes, i.e., "Corrective Action Program" and formal "Safety Policies" (management); and Level 3: attitudes (culture and outcomes). *Safety culture* encompasses all 13 safety culture components developed in the initiative for the enhanced ROP. *Safety management* encompasses only a subset of the safety culture components. For example, the safety-conscious working environment components are aspects of safety culture but not safety management.

NRC Staff Exhibit NRC000055, Enclosure 3, at 2-3.

In the conclusion section of this enclosure, the Staff recommended using the term safety culture in enhancing the ROP because:

> Safety culture encompasses the three levels of Dr. Schein's model of organizational culture (the basis for most safety culture definitions), whereas "safety management" encompasses only the top level (artefacts) and part of the middle level (espoused values). The lowest level, "basic assumptions," can be of greatest concern to the NRC and other regulatory and industry bodies, because often this level encompasses the root causes that may lead to repetitive and far reaching safety problems. However, this level is not addressed by the most common definitions of "safety management" (cf. INSAG-13).

NRC Staff Exhibit NRC000055, Enclosure 3, at 4.

Thus, safety management does not address the basic assumptions of an organization's safety culture (Level 3 of Dr. Schein's model), which can be the most important to the NRC if there are repetitive and significant performance problems.

Q7. Does Mr. Grimes experience in nuclear reactor safety management qualify him to opine on safety culture?

A7(a). (Barnes) No. Mr. Grimes' experience, as documented in his resume and testimony, do not appear to include the full set of knowledge, training, and/or experience one would expect of a safety culture expert. His undergraduate degree is in nuclear engineering. He does not have a graduate-level degree in one of the appropriate social sciences that would provide him with demonstrated knowledge of the theory and research literature applicable to safety culture and related topics, appropriate research methods for studying and analyzing safety culture, the strengths and limitations of the different research and evaluation methods, and the professional ethics of the discipline. His resume does not indicate that he has conducted safety culture assessments, either under supervision or independently.

A7(b). (Cai) No. Safety management generally does not include key aspects of safety culture, as discussed in my response in A6(b). Therefore, the safety management experience Mr. Grimes' describes in his resume and testimony does not qualify him as a safety culture expert. To be qualified as a safety culture expert, Mr. Grimes would have to demonstrate specific training and experience related to understanding and applying those concepts and areas under safety culture that are not covered by safety management (i.e., specific training and experience regarding the basic assumptions level of the organization of Dr. Schein's model and how to assess that level). In addition, he would need specific training and experience related to interpreting the results of safety culture assessments (see response to next question) and how to improve an organization's culture to be viewed as qualified to opine on safety culture.

Q8. Does Mr. Grimes' experience in nuclear reactor safety management qualify him to interpret the results of a safety culture assessment?

A8(a). (Barnes) No. Specifically, he does not appear to have the appropriate education, training, and experience that would provide him with the ability to evaluate the likely validity and reliability of the methods used in a safety culture assessment, the weight to place on different sources and types of information gathered in the assessment, how to best integrate information obtained from disparate sources, and how to arrive at well-founded, defensible conclusions about the generalizability of the results. Mr. Grimes' resume and testimony indicate that he does not have a graduate-level degree in a social science that focuses on human behavior in group settings, supervised experience in conducting either research or practice related to safety culture, or experience conducting safety culture assessments in the field.

A8(b). (Cai) Safety management does not generally encompass the "basic assumptions" of the organization's safety culture, which are not readily observable. (See A6(a) Barnes and A6(b) Cai above). Safety culture assessments typically include techniques and methods, such as interviews, focus groups, and surveys, to gain specific insights into this level of the organization. Mr. Grimes' resume and testimony do not indicate that he has any training or experience with these types of techniques or methods used as part of safety culture assessments. In addition, specific experience with analyzing the results of these methods and techniques and their interpretation in the larger context of an overall organizational assessment (i.e., how to consider and evaluate all the data sources together in context) is needed to make valid interpretations of results from safety culture assessments. Dr. Barnes' response in A5 provides further details on the skills needed for interpretation of results. Again, it does not appear from Mr. Grimes' resume and testimony that he has training or experience evaluating and interpreting the results of safety culture assessments.

Q9. Does Mr. Grimes make any unsupported assertions in his testimony?

A9(a). (Cai) Yes. On pages 19-20 of his testimony (A44), the examples Mr. Grimes cites as indicative of a weak safety culture do not equate to a full safety culture assessment, which is necessary to support an overall conclusion about an organization's safety culture. In a comprehensive safety culture assessment, these examples would be data points that would be evaluated along with many additional sources of information to render a full assessment. Under the ROP, a full safety culture assessment would include evaluation of all 13 safety culture components (see RIS 2006-13 (NRC Staff Exhibit NRC000046) Enclosure at 3), and include all the levels under Schein's model (see discussion in question A6(b)). Mr. Grimes states there are continuing human performance and problem identification and resolution issues at Prairie Island. However, under the ROP the full set of safety culture components also includes several in the safety conscious work environment area, and four additional safety culture components that are not related to one of the cross-cutting areas (i.e., accountability, continuous learning environment, organizational change management, and safety policies). In addition, simply citing these examples does not provide information on the basic assumptions level of Schein's model, which is not readily observable. Data at that level typically need to be gathered through techniques such as surveys, focus groups, and interviews, which are not included in the examples Mr. Grimes cites. See attachment 02 to Inspection Procedure 95003, "Guidance for Conducting an Independent NRC Safety Culture Assessment," (NRC Staff Exhibit NRC000043) for a detailed description regarding what constitutes a safety culture assessment.

A9(b). (Klett) Yes. Mr. Grimes' July 30, 2010 testimony appears to make the following unsupported assertion:

> A38: "The messages also include a description of a regulatory perception: 'Our behavior does not demonstrate that we understand the significance or the uniqueness of nuclear power, nor do we appear to respect the power of the reactor.'"

This statement is taken from a PINGP-generated document, "Required Briefing by Department Managers" (PIIC Exhibit 19). I am not aware of any NRC documents that state that the NRC has this "regulatory perception" of this licensee.

Q10. Does Mr. Grimes make any incomplete statements about the ROP overall?

A10. (Klett) Yes. The following statements from Mr. Grimes about the ROP are incomplete, incorrect, or have incorrect implications:

1. A25: "NRC inspection findings are classified by color designations that indicate the severity of the inspection concern. "Green" designates acceptable performance. However, "white", "yellow", and "red" findings indicate more serious safety problems."

Inspection Manual Chapter (IMC) 0308, Reactor Oversight Process Basis Document (IMC 0308) (NRC Staff Exhibit NRC000008), defines the meaning and thresholds of acceptable performance, as noted below.

- The licensee response band is characterized by acceptable performance in which cornerstone objectives are fully met; nominal risk with nominal deviation from expected performance. This performance band is designated as the Green band. Performance problems would not be of sufficient significance that escalated NRC engagement would occur. Licensees would have maximum flexibility to "manage" corrective action initiatives.

- The increased regulatory response band would be entered when licensee performance is outside the normal performance range, but would still represent an acceptable level of performance. This performance band is designated as the White band. Cornerstone objectives met with minimal reduction in safety margin; outside bounds of nominal performance; within Technical Specification Limits. Degradation in performance in this band is typified by changes in risk of up to 10^{-5} ΔCDF or 10^{-6} ΔLERF associated with either PIs or inspection findings. The CDF and LERF threshold characteristics were selected to be consistent with RG 1.174 applications.

- The required regulatory response band involves a decline in licensee performance that is still acceptable with cornerstone objectives met, but with significant reduction in safety margin; Technical Specification limits reached or exceeded. This performance band is also designated as the Yellow band. Degradation in performance in this band is typified by changes in risk of up to 10^{-4} ΔCDF or 10^{-5} ΔLERF associated with either PIs or inspection findings. These threshold characteristics and required regulatory response are also selected to be consistent with risk-informed regulatory applications and mandatory actions for regulatory compliance.

- The Red band is typified by changes in performance that are indicative of changes in risk greater than 10^{-4} ΔCDF or 10^{-5} ΔLERF associated with either PIs or inspection findings. Plant performance represents an unacceptable loss of safety margin. It should be noted that should licensee's performance result in a PI reaching the Red Band, margin would still exist before an undue risk to public health and safety would be presented.

In short, neither a White Finding, nor a plant transitioning to the regulatory response Column indicates that a licensee's performance level is unacceptable in accordance with IMC 0308 (NRC Staff Exhibit NRC000008).

2. A28. "For plants in the Regulatory Response column, the NRC conducts additional inspections beyond the normal inspection program, and takes other actions, to focus on potential safety problems (approximately 10 to 20% of all operating reactors are in the Regulatory Response column)."

Regarding the statement, "[T]he NRC … takes other actions, to focus on potential safety problems […]," it is not clear what Mr. Grimes considers to be "potential safety problems" or what other actions the NRC would take in response to these. The NRC's actions are described in IMC 0305, Operating Reactor Assessment Program (0305) (NRC Staff Exhibit NRC000011) and supplemental inspection procedures.

3. A40: "The CAP is the program the applicant relies upon to effectively accomplish the Problem Identification and Resolution expectations of the Reactor Oversight Process."

Not all actions necessarily need to be handled within the licensee's corrective action program under 10 CFR 50, Appendix B, Criterion XVI. It may be more appropriate for some issues that are not conditions adverse to quality to be tracked to resolution through an alternate licensee program such as an employee concerns program. The NRC inspects these programs in addition to the corrective action program. For example, issues identified from a safety culture assessment may be resolved in a licensee's employee concerns program rather than its corrective action program.

From IP 71152 (NRC Staff Exhibit NRC000051):

> If the licensee conducted any periodic self-initiated assessments of safety culture during the review period, this assessment shall be included along with other non-safety culture self-assessments selected to review. If the licensee performed several assessments that collectively addressed safety culture issues, then those assessments combined should be considered as one assessment. [C2] Inspectors should review the adequacy of the licensee's evaluation and actions to address the issues identified by the safety culture assessment. Not all actions necessarily need to be handled within the licensee's corrective action program under

> 10 CFR 50, Appendix B, Criterion XVI. It may be more appropriate for some issues that are not conditions adverse to quality to be tracked to resolution through an alternate licensee program such as an employee concerns program. The inspectors review should focus mainly on the licensee's response to the assessment results or actions taken to address identified issues instead of the assessment methodology or an evaluation of the assessment's adequacy.

Q11. Does Mr. Grimes make any incomplete or incorrect statements about the ROP's oversight of safety culture?

A11(a). (Keefe & Klett) Yes. The following statement from Mr. Grimes about the ROP's oversight of safety culture may be incorrect or have incorrect implications:

> A26. "The components of safety culture are directly related to the crosscutting areas of human performance, a safety conscious work environment, and problem identification and resolution."

The safety culture components are comprised of 13 components, 9 of which are cross-cutting. A safety culture assessment, which IMC 0305 (NRC Staff Exhibit NRC000011) defines as a comprehensive evaluation of the assembly of characteristics and attitudes related to all of the safety culture components, would look at all 13 safety culture components. Nine of the thirteen safety culture components are linked to the cross-cutting areas; however, there are four additional components that are not located under the cross-cutting areas. These additional four components, which are accountability, continuous learning environment, organizational change management, and safety policies, are reviewed only during supplemental inspections. Regulatory Issue Summary 2006-13, "Information on the Changes made to the Reactor Oversight Process to More Fully Address Safety Culture," (NRC Staff Exhibit NRC000046) provides the following explanation:

> Safety culture components 1–9, termed 'cross-cutting components,' are aligned with the three cross-cutting areas (i.e., human performance, problem identification and resolution, and safety conscious work environment) and replace the existing cross-cutting subcategories or bins. However, the supplemental inspection program applies to all 13 safety culture components. This distinction was made because of the following:
>
> - The nine cross-cutting components are currently readily accessible

through baseline inspection procedures, while the last four safety culture components listed above (i.e., accountability, continuous learning environment, organizational change management, and safety policies) are not.

• Each of the nine cross-cutting components is closely aligned with the cross-cutting area with which it is associated, while components 10–13 listed above are not closely aligned with a cross-cutting area.

A11(b). (Klett) Yes. The following statements from Mr. Grimes about the ROP's oversight of safety culture may be incorrect or have incorrect implications:

1. Q34: "How do the performance issues associated with all of the recent White findings relate to safety culture?" A34: "These human performance crosscutting issues are one aspect of the NRC's assessment of the "safety culture" at an operating reactor. In addition to the human performance crosscutting issue identified in relation to the PINGP White findings […]."

Substantive cross-cutting issues (SCCIs) do not constitute safety culture assessments, as defined in IMC 0305 (NRC Staff Exhibit NRC000011). If a licensee has difficulty in correcting a long-standing SCCI, the NRC may request the licensee to perform a safety culture assessment to determine if any weaknesses in safety culture contributed to the difficulty in correcting the long-standing SCCI.

2. A44: "The failure of the applicant to correct the potential damage to the containment integrity resulting from the refueling cavity leaks, the safety culture weaknesses associated with the causal factors described in Information Notice 2009-11, the series of White findings associated with one or both of the PINGP units, the identification of substantive crosscutting issues in the area of human performance, the serious concerns identified by NRC inspectors with the applicant's corrective action program, and failure to effectively manage the plant design and effectively resolve potentially the safety-significant flooding issues identified 20 years ago, are all indicative of a weak safety culture at PINGP.

Again, inspection findings, substantive cross-cutting issues, and NRC inspection reports do not constitute safety culture assessments, as defined in IMC 0305 (NRC Staff Exhibit NRC000011). If a licensee has safety-significant inspection findings and/or a long-standing substantive cross-cutting issue (SCCI), the NRC may request a licensee to perform a safety culture assessment to determine if any weaknesses in safety culture contributed to the performance issues or the difficulty in correcting a long-standing SCCI.

Q12. Does Mr. Grimes make any unsupported conclusions about the current status of safety culture at PINGP?

A12. (Cai) Yes. On page 20 of his testimony (A44), Mr. Grimes cites a list of examples he claims are indicative of a weak safety culture at PINGP. However, there are no agreed-upon or empirically validated measures or thresholds, either by the NRC or in the nuclear industry, that provide indication of the strength or health of an organization's safety culture (i.e., for determining what constitutes a "weak" or "strong" or "healthy" safety culture). In other words, there are no empirically validated measures that, in and of themselves, provide an indication of the strength or health of an organization's safety culture. Because of this, safety culture assessment results are often compared with assessment results from other organizations of similar size, structure, and/or function, using the same methodology. Results from safety culture assessments can also be trended over time to provide indication of improvements or declines in the organization's safety culture. Mr. Grimes' assertions of a weak safety culture at PINGP do not have a basis because there is no commonly accepted measure or threshold of what defines a "weak" safety culture, and he does not offer any comparison or trending information to provide support for his assertion.

Q13. Citing the Commission's draft policy statement on safety culture (NRC Staff Exhibit NRC000056), Mr. Grimes asserts in A34 that "a fundamental characteristic of an effective safety culture is that '[t]he organization ensures that issues potentially impacting safety or security are promptly identified, fully evaluated, and promptly addressed and corrected….'" Do you agree?

A13. (Cai) This is Mr. Grimes' opinion. In the draft policy statement, the actual reference regarding the characteristics is: "The following characteristics that are indicative of a positive safety culture" (NRC Staff Exhibit NRC000056 at 57528). Further, in the implementation section, the draft policy statement states,

> This policy statement describes areas important to safety culture, but it does not address how the nuclear industry, the Agreement States, and the NRC should establish and maintain a positive safety culture in their organizations. The nuclear industry, the Agreement States, and the NRC differ in their size and complexity, infrastructure, and organizational frameworks. Therefore, a single approach for establishing and maintaining a positive safety culture is not possible…The Commission expects…all organizations [to] consider and foster the safety culture characteristics (commensurate with the safety and security significance of activities and the nature and complexity of their organization and functions) in carrying out their day-to-day work activities and decisions

Id. at 57529). The characteristics of a positive safety culture described in the draft policy statement are not standards which licensee organizations can be specifically evaluated against, and the implementation of the expectations in the draft policy statement is not in the scope of the policy statement and can vary depending on the type of organization. Therefore, Mr. Grimes' reference to the policy statement in his response does not provide any specific support for his assertion regarding the applicant's "ability to resolve other potentially risk-significant deficiencies."

Q14. In A45 of his testimony, Mr. Grimes asserts that the NRC should direct the applicant to conduct a third party safety culture assessment per IMC 0305 and determine what corrective actions are necessary, in order to provide reasonable assurance that the applicant will manage the effects of aging. Does IMC 0305 contain procedures for a third party safety culture assessment?

A14. (Cai) No. Currently, there is no standard methodology for conducting a third-party safety culture assessment, although the industry is working on developing one (see A21 (Barnes)). Second, the reference in IMC 0305 regarding third-party assessments is only with regard to the composition of the team; it does not specify what methodology and techniques the applicant should undertake. In addition, as I stated in A12, there are no agreed-upon or empirically validated measures or thresholds, either by the NRC or in the nuclear industry, that provide indication of the strength or health of an organization's safety culture (i.e., for determining what constitutes a "weak" or "strong" or "positive" safety culture).

Q15. Do you agree, as Mr. Grimes asserts in A45 of his testimony, that a third party assessment would determine what corrective actions are necessary in order to provide reasonable assurance that the applicant will manage the effects of aging?

A15. (Barnes) I disagree with Mr. Grimes' assertion. The process of conducting a safety culture assessment is distinct from developing corrective actions. Safety culture assessments identify problems and opportunities for improvement in an organization's safety culture, but may not provide sufficient information and analysis to develop corrective actions. After receiving results of the assessments that are typically conducted in the nuclear power industry, licensees often must gather additional information to understand the bases for the perceptions and attitudes identified in the assessment. This additional information is then used to design appropriate interventions. Some organizations request an assessment team to thoroughly investigate any weaknesses identified and offer recommended corrective actions. However, the scope of an assessment and the design of appropriate corrective actions are unrelated to assessment team composition (i.e., one that meets the IMC 0305 definition of a third party assessment).

Q16. Mr. Grimes asserts that "simple follow-up inspections" are inadequate to ensure that "culture corrective measures are both effective and sustainable." Does the ROP provide for more than simple follow-up inspections?

A16(a). (Keefe) The ROP provides opportunities to gather insights into safety culture during routine baseline inspections. As described in our original testimony, if a licensee has safety significant issues, the ROP provides for more in-depth review of safety culture through the supplemental inspection procedures (IP 95001, 95002, and 95003 (NRC Staff Exhibits NRC000013, NRC000042 and NRC000026). PINGP has not reached the performance threshold at which the NRC would request an additional safety culture assessment at this time.

A16(b). (Klett) Yes. First, the NRC uses IP 71152 (NRC Staff Exhibit NRC000051) to follow-up on substantive cross-cutting issues (SCCIs). This procedure allows for continuous and periodic follow-up. For SCCIs, the NRC is evaluating a licensee's corrective actions for the SCCIs; therefore, IP 71152 is adequate. If the NRC requests the licensee to perform a safety culture assessment in response to a long-standing SCCI, the inspectors may use portions of the guidance in IP 95003.02 (NRC Staff Exhibit NRC000043) to review the adequacy of the licensee's response to a safety culture assessment.

The NRC may request that PINGP perform an independent safety culture assessment if the SCCI mentioned in the NRC's March 3, 2010, assessment letter to PINGP (Applicant Exhibit NSP000032) remains open after the 2010 mid-cycle performance review.

If PINGP were to transition to the Degraded Cornerstone Column (Column 3) of the ROP Action Matrix, the NRC would perform an IP 95002 (NRC Staff Exhibit NRC000042) inspection. As part of that inspection, the inspectors would perform a focused inspection to independently determine that the licensee's root cause evaluation appropriately considered whether any safety culture component caused or significantly contributed to any risk-significant performance issue. If a weakness in any safety culture component did cause or significantly contributed to such an issue, and the licensee's evaluation did not recognize that cause or contribution, the NRC may request the licensee to perform an independent safety culture assessment.

If PINGP were to transition to the Multiple/Repetitive Degraded Cornerstone Column (Column 4) of the ROP Action Matrix, the NRC would request that the licensee have a third-party safety culture assessment performed. If the licensee refuses, the NRC would perform the assessment in accordance with IP 95003.02 (NRC Staff Exhibit NRC000043). The NRC follows-up on the licensee's third party assessment using IP 95003 (NRC Staff Exhibit NRC000026).

Thus, the ROP includes more than simple follow-up inspections to ensure that licensees are addressing safety culture issues.

Q17. Do you agree, as PIIC asserts, that PINGP's treatment of the refueling cavity issue is a culminating symptom of a weak safety culture?

A17. (Keefe) I would disagree with the assertion that the refueling cavity leakage issue can be linked to a weak safety culture. The NRC never made a finding relating to refueling cavity leakage and containment integrity. Because there has been no finding, the NRC never determined that the cause of refueling cavity leakage was associated with a cross-cutting aspect. Furthermore, Steven C. Skoyen testified that PINGP has a good understanding of the source of the leak, and inspections on the areas of containment impacted by the leak have found no damage to the containment vessel. Jack Giessner testified that when the refueling cavity leak was identified, the NRC had structural experts review the issue, to determine whether there was an immediate safety issue. The assessment concluded there was not. The NRC's most recent inspection findings on this issue are discussed in NRC Inspection Report No. 500282/2010003; 05000306/2010003 (July 26, 2010) (NRC Staff Exhibit No. NRC000018). This report states that the NRC did not identify any findings of significance, and notes that previous evaluations have not revealed any degradation of the containment pressure vessel, concrete, or rebar due to the refueling cavity leakage. Based on the testimony provided, PINGP has taken steps to understand the historical leakage issue and to mitigate containment degradation. Thus, I disagree with PIIC's assertion that PINGP's treatment of refueling cavity leakage is a symptom of weak safety culture.

Q18. Do you agree with Mr. Grimes' description of the NRC's "most recent annual assessment letter"?

A18. (Klett) No. In A44 Mr. Grimes states:

> "There is a pattern of cultural performance issues revealed by the continuing human performance (HP) and problem identification and resolution (P&IR) issues at Prairie Island that go too deep to be addresses by a simple follow-up inspection. As described in the NRC's most recent Annual Assessment Letter, the Applicant must demonstrate that the cultural corrective measures are both effective and sustainable."

Mr. Grimes does not accurately characterize the content of the NRC's annual assessment letter to Prairie Island dated March 3, 2010 (Applicant Exhibit NSP000032). This letter does not state that the licensee must demonstrate that the "cultural corrective measures" are both effective and sustainable. The NRC does not characterize the SCCI as a "cultural performance issue." The NRC's letter states:

> You discussed your actions to address the ongoing HP SCCI during a December 1, 2009, public meeting and some improvement has been observed. However, these actions have not yet proven effective in mitigating the cross-cutting themes. Therefore, the NRC has a concern with your progress in addressing this cross-cutting area and has concluded that the SCCI will remain open until all cross-cutting themes have been cleared in the area of HP. To clear each of the themes, we will assess whether your corrective actions have resulted in a positive sustainable improvement in the area and will consider the number of findings in the theme.

Applicant Exhibit NSP000032.

When this letter was written, the SCCI was not considered to be long-standing. The NRC may request the licensee to perform a safety culture assessment if the SCCI is open for 18 months.

Q19. Have you reviewed Prairie Island Station Nuclear Safety Culture Assessment (Applicant Exhibit NSP000057)?

A19. (Barnes) Yes.

Q20. What type of safety culture assessment is this, i.e., self, independent, or third-party?

A20. (Barnes) This safety culture assessment is a peer- or self-assessment. It does not meet the NRC's definition of an independent or a third-party assessment because the assessment team included site employees (referred to as "peer" team members in the assessment report) who had input into the assessment team's conclusions. Section 04.17 of NRC Inspection Manual Chapter 0305 defines independent and third-party safety culture assessments:

> A licensee independent safety culture assessment is performed by qualified individuals that have no direct authority and have not been responsible for any of the areas being evaluated (for example, staff from another of the licensee's

> facilities, or corporate staff who have no direct authority or direct responsibility for the areas being evaluated). A licensee third-party safety culture assessment is performed by qualified individuals who are not members of the licensee's organization or utility operators of the plant (licensee team liaison and support activities are not team membership).

NRCStaff Exhibit NRC000011.

Q21. What is your opinion of the methods used to develop the Assessment?

A21. (Barnes) The safety culture assessment report does not include a detailed description of the methods used to conduct this assessment and I did not observe the process or have access to the raw data from the assessment. Therefore, there are limitations on the bases for my opinion. The methods used appear to be the ones currently being proposed by the industry. As such, I am familiar with the methods used.

When reviewing a safety culture assessment, the NRC Staff looks for methods that support the reliability and validity of the results and conclusions. Reliability is measured and evaluated in several ways, but the central concern is whether the same results would be obtained if different assessors were involved, if different methods were used, and if different representatives of the organization being assessed (e.g., interviewees) participated in the assessment. Validity is also measured and evaluated in a number of ways, but the central concern for validity is whether a method or set of methods actually measures the construct(s) of interest and nothing else. A method cannot be valid if it is not also reliable.

Although the report does not provide a detailed description of the methods used in the assessment, it states on page 2, that "…this assessment meets the guidance provided in NEI 09-07, *Fostering A Strong Nuclear Safety Culture* [NEI 09-07, Industry Guideline, *Fostering a Strong Nuclear Safety Culture*, Rev. A., Nuclear Energy Institute]." The NRC Staff reviewed NEI 09-07, Rev 0, and, at NEI's invitation, observed three pilot assessments performed in accordance with that version. The NRC Staff provided its observations on the assessment process to NEI at a public meeting held on February 24, 2010. *See* NRC Staff Exhibit NRC000057. With respect to its review and observations of the three pilot assessments conducted under Rev 0 of NEI 09-07, the Staff had a number of substantive concerns with (1) the interview process; (2) the ratings team members assigned to interview and observation results; (3) the validity and reliability of the survey used; and (4) the method for aggregating information to reach conclusions.

Staff concerns with the interviews we observed included that: (1) it appeared that some assessment team members had not been trained to conduct interviews in a manner that ensures rapport with interviewees and avoids biasing their responses; (2) the presence of site peers, including some supervisors, in the interview sessions, appeared to cause discomfort for some interviewees; (3) the interview questions were complex and appeared to be difficult for interviewees to understand, with the result that, in order to communicate, interviewers reworded the questions, but did so inconsistently and sometimes changed the apparent meaning of the questions; and (4) the selection of questions for the interviews appeared to be ad hoc. These issues led the Staff to question the reliability of the interview process, and therefore its validity.

The Staff's concern about the method used to rate interview and observation results was that team members did not appear to agree on whether an interview response or observation represented a positive, neutral, or negative indication of safety culture. The Staff also noted that the different teams at the three different pilot sites rated the interview responses and observations differently. These inconsistencies also led the Staff to question the reliability and, hence, the validity of the assessment results.

The Staff's concern with the survey used included: (1) the psychometric properties of the survey had not been established and (2) no industry norms were available against which to compare one site's responses to others across the industry. Without analyses demonstrating the psychometric properties of the survey, its validity and reliability are also questionable. The absence of norms precludes an objective, comparative evaluation of any conclusions regarding such safety culture descriptors as "strong" or "positive."

Finally, the Staff was concerned with the method used to aggregate the data points obtained from the assessments. The NSCA method aggregates data related to each INPO principle and attribute to provide a site-wide summary conclusion for each. However, large organizations, such as nuclear power plants, include many local subcultures that can exist at various levels, ranging in scope from small work units (e.g., a first-line supervisor and 2 workers in a team) to entire departments (e.g., Operations). As a result of aggregating data by principle or attribute, subculture differences are not identifiable. Therefore, the scope of claimed site-wide safety culture strengths or weaknesses is unknown. For example, perhaps Operations at a site is experiencing significant safety conscious work environment (SCWE) issues and is very unhappy with their managers, trainers, and one another; but all other site departments are doing well. Because the Operations' results will be aggregated with those of other departments by principle or attribute, the severe problems in this hypothetical Operations department will be masked and the site's overall results will be lowered, when in fact it's a local problem that may not be representative of safety culture at the site overall.

During a July 28, 2010 public meeting, NEI stated that it is working to address the concerns the Staff expressed at the February 24, 2010, public meeting. *See* NRC Staff Exhibit NRC000058. However, NEI has not yet submitted, so the Staff has not reviewed, Rev A of NEI 09-07. Further, the NRC Staff did not observe the safety culture assessment conducted at PINGP. It is possible that the interview and team rating processes during the assessment at PINGP were improved, compared to those the Staff observed at the pilot plants. The PINGP June, 2010 report (Applicant Exhibit NSP000057) provides no evidence that the survey administered is different from that used at the pilot plants, however, and does not present information regarding the survey's psychometric properties or normative data. In addition, the assessment results in the report are presented by principle and attribute, again with no information about possible subculture effects. In this specific assessment, I am also concerned that the assessment team's observations of workforce behaviors were limited to observation of six meetings. Six meetings is a small sample from which to draw conclusions about meeting behavior. In addition, meeting behavior is a limited sample of the many types of workforce behaviors that may be indicative of safety culture. The assessment did not include observations of actual work being performed, for example, by maintenance or health physics personnel on the shop floor or activities in the control room, in work planning, in engineering, etc. Therefore, the behavioral observation data collected during the assessment may not be representative of workforce behavior at the site.

Despite these concerns, this assessment had one strength compared to those performed at the pilot plants. That is, the survey response rate was higher (88% at PINGP) than at the pilot plants. There is no evidence presented to support the report's assertion that this response rate indicates the workforce's strong commitment to safety culture at the site, and there are many other possible explanations for this high response rate. But, the high response rate demonstrates that the large majority of the workforce had and took the opportunity to respond. This suggests that the survey provided comprehensive information about perceptions and attitudes among the site workforce. Even so, the report does not offer information about the characteristics of those who did not participate in the survey. For example, did the 12% of non-respondents consist of a particular workgroup, such as security which often has a low response rate on safety culture surveys, or was that 12% apparently randomly distributed among work groups?

Q22. Is the June, 2010 Assessment (Applicant Exhibit NSP000057) probative of the current status of safety culture at PINGP? Why or why not?

A22. (Barnes) I cannot offer a conclusion about whether the assessment is probative or not for the reasons described above. I am comfortable opining that the assessment likely gathered a reasonable amount of information about the workforce's perceptions of the current state of safety culture at the site, but I do not have confidence in the assessment's conclusions regarding the "strength" of the site's safety culture.

Q23. Can you make an overall conclusion about the current status of safety culture at PINGP based upon the June, 2010 Assessment?

A23. (Barnes) No. I would need additional information that the assessment report does not provide to offer an overall conclusion about the current status of safety culture at PINGP. For example, I would want to review the actual survey data and write-in comments, notes from the interviews and the questions asked; to have observed the team's rating processes; have access to normative data regarding how PINGP's results compare to other sites'; and have access to information collected from the different workgroups at the site, and other information, such as a wider range and larger sample of behavioral observations.

Q24. Can you make any predictions about safety culture during the proposed period of extended operation based on the Assessment (Applicant Exhibit NSP000057)?

A24. (Barnes) No. To the extent that I understand the approach and methods used, this assessment evaluated PINGP's current safety climate (cf. my July 30, 2010 testimony at A4 regarding the differences between safety culture and safety climate). The assessment does not appear to have collected information about any deep-seated, underlying assumptions embedded in the safety culture at PINGP that may potentially be impervious to change over the next 20 years. Further, the assessment appropriately does not attempt to predict into the future by assuming that there will not be any external or internal influences to drive change – which is unlikely, as I've previously testified. This assessment provides a picture of the safety climate at PINGP at the time the assessment was done. The assessment's relevance to the site's future safety culture over the period of extended operation is minimal.

Q25. What is your expert opinion of the overall current status of safety culture at PINGP?

A25. (Barnes) The record is inadequate to support an overall opinion on safety culture at PINGP. Based on the documents I have reviewed, I am comfortable offering an opinion that the site appears to face some safety culture challenges. For example, like the industry as a whole, the record indicates that PINGP's senior workforce is retiring, so it is faced with hiring challenges in a depleted nationwide candidate pool and knowledge transfer issues as new hires replace senior staff. Retirements and an influx of new hires, however, is also an opportunity for the site to overcome an attitude described in the March, 2010, Management and Review Safety Committee meeting report (PIIC Exhibit 21) that use of human performance tools, such as assiduous placekeeping and step-by-step execution of procedures, is not necessary for "experienced" personnel. As another example, communication of the bases for disciplinary decisions to hold individuals accountable for using human performance tools and consistency in those decisions appears to be a continuing challenge (Applicant Exhibit NSP000057 at 10). It is a rare organization that does not experience challenges to maintaining and improving its safety culture as it reacts to external and internal influences for change. It unclear from the record, however, that PINGP's safety culture overall is deficient.

Q26. In A45, Mr. Grimes concludes that, in order for the NRC to have reasonable assurance that PINGP will manage the effects of aging during the period of extended operation, the NRC should direct PINGP to conduct a third-party assessment of safety culture and that the NRC should address corrective actions by PINGP before the NRC issues a renewed license. Do you agree?

A26. (Barnes) I disagree. Although he provides no rationale for his conclusions, Mr Grimes' recommendation appears to be based on his view that knowledge of PINGP's current safety culture will be predictive of PINGP's safety culture over the period of extended operation, which is an erroneous assertion I addressed in my previous testimony. Further, Mr. Grimes also appears not to recognize that corrective actions taken now or in the near-future for identified safety culture weaknesses similarly will not ensure that PINGP's safety culture will remain adequate over the period of extended operation. In my opinion, the on-going, day-to-day inspections and monitoring of PINGP safety performance that are required under the ROP are the most appropriate approach to ensuring that PINGP continues to address safety culture challenges as they arise and implements its aging management program effectively.