



Union of Concerned Scientists

Citizens and Scientists for Environmental Solutions

MRE
TS
LLJ
ADAMS

August 18, 2005

Michael R. Johnson, Director
Office of Enforcement
U.S. Nuclear Regulatory Commission
Washington, DC 20555-0001

SUBJECT: SAFETY CULTURE WITHIN THE REACTOR OVERSIGHT PROCESS

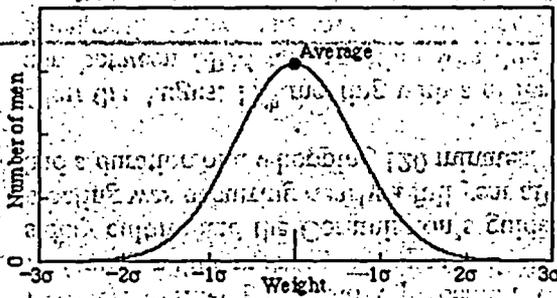
Dear Mr. Johnson:

I am writing to you in your capacity as Chairman of the Safety Culture Steering Committee following the August 17th public meeting on the subject.

I confess to having a negative perception when Ms. Isabelle Shoefeld first contacted me about the then upcoming public meeting. This perception stemmed from my belief that the NRC staff must not be taking safety culture and the Commission's guidance in SRM-04-0111 seriously given the fact that the public meeting was occurring nearly a full year after the August 30, 2004, date of the SRM and that the meeting had a duration of a whopping 120 minutes.

I left the August 17th meeting with a different perception. It appeared to me that the reason for the year gap between SRM and meeting was due to the staff completing considerable homework rather than because the issue has a low priority. It also appears to me that the duration of the meeting was due to its function as a kick-off session for a planned series of public interactions rather than because that was the minimum amount of time needed to check off some box on a form. I hope this revised perception is the correct one.

This history is provided primarily to provide foundation for my comment about the hazards with using numbers without proper context. Armed with numbers (year gap and short meeting duration) but lacking underlying context, I leapt to the not-unreasonable conclusion that the NRC staff was not taking safety culture seriously. I agree with Jerry Roberts and several of the industry representatives that, without the proper underlying context, many of the numbers filled in on the safety culture attributes table could provide an inaccurate perception.



I am by no means a safety culture expert. But having listened to many such experts over the past decade, it is clear to me that safety culture is difficult to measure. The accuracy of safety culture assessments is inversely related to the bell curve. In other words, safety culture assessments can accurately identify very good and very bad conditions, but struggle to accurately characterize the majority of conditions that fall in the middle somewhere.

Safety culture assessments are much more accurate when characterizing trends than when making "snapshot" determinations. Thus, an assessment may be imprecise when attempting to define the absolute soundness of safety culture at that moment yet be very insightful when defining that moment's safety culture relative to the culture that existed six months or a year ago.

With this understanding about safety culture assessments, I am more than a little leery about equipping all NRC inspectors with "safety culture sonar." Given the large number of tasks performed at nuclear power plants and the complexity level of many of these tasks, there will always – ALWAYS – be data that could suggest safety culture problems. Rather than sending an army of NRC inspectors out on safety culture scavenger hunts with the very likely outcome of having them dredge up this data, relevant or not as signs of bad safety culture, it would seem far, far better for the inspectors to focus on assessing performance levels as accurately as possible and then having safety culture experts determine whether bad safety culture could be a cause for declining performance. It seems very unlikely to highly improbable that, for example, an inspector conducting a radiation protection inspection will be able to accurately assess safety culture at that "snapshot." It seems equally unlikely/improbable that any Branch Chief will be able to take a safety culture "snapshot" from a radiation protection inspector along with a "snapshot" from a fire protection inspection and get the correct impression of the safety culture trend. If so, they would be called Branch Wizards instead of Branch Chiefs.

Instead, NRC inspectors could conduct their radiation protection, fire protection, and numerous other inspections aimed at assessing performance in areas of their expertise. When they detect declining performance, the safety culture experts could be called in. The safety culture inspectors' initial assessment would simply be a baseline for subsequent trend analysis as the NRC oversees the licensee's efforts to stop the performance decline and restore it. Because safety culture trend analysis seems far more reliable than instantaneous "snapshots," this approach should enable the NRC to ensure that both the performance deficiencies and their causes are corrected.

As you noted at the onset of the August 17th public meeting, the reactor oversight process was developed based on an assumption that cross-cutting performance issues, including bad safety culture, would manifest themselves in the performance indicators and/or inspection findings. Some cite Davis-Besse and Salem/Hope Creek as *prima facie* evidence this assumption is invalid. I am not persuaded that is the case. I am convinced that the reactor oversight process can and should do a better job of handling safety culture, but I am far from convinced that tasking all NRC inspectors with performing safety culture assessments – even of limited scope – is the right approach.

The performance indicators and inspection findings did a deplorable job of characterizing the serious safety problems caused at Davis-Besse by the bad safety culture. Recall that, after the fact, the NRC issued a RED finding for the reactor coolant system leakage leading to reactor vessel head wastage, a YELLOW finding for longstanding containment sump deficiencies, a WHITE finding for deficient high pressure injection pump design dating back to original construction in the late 1970s, and two other WHITE findings for radiation protection violations. Apparently, bad safety culture played a role in these unwanted conditions being present at Davis-Besse for so long. Ample evidence of these many serious problems existed, but all the signs were totally missed by the performance indicators and inspections. The proper NRC solution to this situation is not, repeat NOT, to develop an army of safety culture inspectors. The proper solution is to enable its existing army of inspectors with the means to have identified one or more of the RED, YELLOW, or WHITE findings sooner and then having a squad of safety culture inspectors come in to determine why that first thing went non-GREEN.

The performance indicators did an equally deplorable job of characterizing conditions at Salem and Hope Creek, but the inspection findings did a much better job. One can argue that, absent external pressure from whistleblower Dr. Kymn Harvin, these inspection findings might still not have identified the underlying bad safety culture at this troubled site. In any case, it is apparent that the NRC inspections did not detect declining performance soon enough, based on results from an independent team's appraisal that rated 73 of 90 areas "less than competent" at Salem / Hope Creek.

In my view, the question to be answered is: **What triggers safety culture monitoring at a site?**

Safety culture inspections could be made part of the baseline inspection program, but that would very likely be ineffective. The baseline inspection program does not have infinite resources. Carving out xx inspection hours for safety culture inspections means there will be xx fewer hours spent inspecting other baseline areas. As those other baseline inspections get even shallower than they were at Davis-Besse, the likelihood of missing existing safety problems increases. Unless that probability is matched or exceeded by the likelihood of having xx inspection hours detect a bad safety culture when one exists, the shift will be counter-productive. I see no evidence for suspecting yet alone believing that NRC inspectors given minimal training will be able to detect anything other than a really good or really bad safety culture – the two ends of the bell curve. For the vast majority of cases falling in the middle, they will be unable to detect with any reasonable certainty the actual safety culture. When they under-rate safety culture, the “unnecessary burden” objective suffers. When they over-rate safety culture, the “maintain safety” objective suffers. The likelihood of avoiding both of these outcomes is so small as to strongly suggest that the exercise simply be avoided.

Rather than set NRC inspectors up for almost certain failure, it would be better to continue to apply the xx inspection hours evaluating performance on the baseline program topics. By not diluting this inspection effort, the odds of detecting an existing performance deficiency will not decrease. Each time the baseline inspection program identifies a performance problem, the NRC’s supplement inspection effort could assess the safety culture component.¹

In their SRM for SECY-94-0111, the Commissioners directed the staff to make a formal determination of whether bad safety culture is responsible for a plant being in the Degraded Cornerstone Column (i.e., column three) of the ROP Action Matrix. In addition to responding at that time, it might also be prudent to perform a safety culture screening for greater-than-GREEN performance indicators and/or inspection findings. This screening would not entail an entire squad of safety culture inspectors needed to render a formal determination about safety culture but would involve a safety culture expert examining the facts surrounding the indicator/finding and see if there’s a compelling case for bad safety culture being a potential cause. This screening might be conducted using a checklist such as that developed by the NRC staff (i.e., the safety culture attributes table) coupled with input from the NRC Resident Inspectors who could provide the underlying context for the attributes. If that screening identified bad safety culture as a potential cause, the NRC’s supplemental inspections could include safety culture experts to probe deeper.

This approach seems practical. If the NRC conducted a safety culture assessment absent some initial performance deficiency, it would lead to nothing without the concurrent revelation of a performance deficiency. For example, had the NRC safety culture team trekked to Davis-Besse in fall 2001 and found the exact same safety culture it found in fall 2003, that finding would have produced no discernible change in the NRC’s regulatory posture unless one or more of the associated symptoms of that bad safety culture (i.e., the football-sized hole in the reactor vessel head or the HPI pump that would not have worked during an accident) had been identified, too. Absent identification of one or more associated symptom, the Significance Determination Process could never be greater-than-GREEN no matter how bad the safety culture assessment. Thus, contrary to the fanciful notion that identification of a bad safety culture could be a leading indicator, it cannot possible be so.

My answer to the question of what triggers safety culture monitoring at a site: Each and every greater-than-GREEN performance indicator and inspection finding initiates a screening by a safety culture expert to see if a bad safety culture is a contributing cause.

¹ Theoretically, each time a performance indicator identifies a performance problem, there would also be an opportunity to assess the safety culture component. But the nuclear industry successfully neutered and undermined the efficacy of the performance indicators since the inception of the ROP that it’ll be a cold day in h*ll before the next non-GREEN performance indicator occurs.

This concept resembles the screening process employed by licensees for changes, tests, and experiments conducted pursuant to 10 CFR 50.59. Virtually all changes, tests, and experiments are screened to see if an unreviewed safety question could potentially be raised. If so, a more elaborate safety evaluation is conducted to formally determine if an unreviewed safety question exists. If so, the change, test, or experiment is not performed until approval from the NRC is received. Likewise, all greater-than-GREEN performance indicators and inspection findings should be screened to see if a deficient safety culture could be involved. If so, a more elaborate assessment should be conducted to determine if a deficient safety culture exists. If so, the licensee should be politely asked to do something about it.

The NRC also needs to integrate its efforts on safety culture with its longstanding process of generic communications. For some unexplained reason, the NRC has not issued generic communications on safety culture issues. For example, when Millstone had its problems in the mid 1990s, the NRC issued Information Notice 96-17, "Reactor Operation Inconsistent with the Updated Final Safety Analysis Report," but issued nothing about the Commission's order in October 1996 requiring independent oversight of the employee concerns program. After Davis-Besse's problem was discovered in March 2002, the NRC quickly issued Information Notice 2002-11, "Recent Experience with Degradation of Reactor Pressure Vessel Head." But after FirstEnergy told the NRC in August 2002 that it placed production ahead of safety, the NRC issued no generic communication. Licensees have developed extensive operating experience programs that formally review NRC generic communications and factor relevant information into procedures and training. But the NRC has to actually issue a generic communication for these licensee operating experience programs to work. They do not take NRC's wishes, desires, or even meeting summaries as input – they need the Information Notice or Bulletin. The reason why the NRC's generic communications process has ignored safety culture problems must be fixed and the NRC simply must issue generic communications on safety culture problems as it has long done for other problems.

Sincerely,



David Lochbaum
Nuclear Safety Engineer