



THE NRC’S REACTOR OVERSIGHT PROCESS: AN ASSESSMENT OF THE FIRST DECADE

The Reactor Oversight Process (ROP),¹ which the Nuclear Regulatory Commission (NRC) instituted to evaluate the safety and security performance of the nation’s 104 nuclear power reactors, recently passed the ten year mark. This issue brief documents UCS’s review of the ROP’s first decade. The NRC appointed UCS to a panel formed in 1999 under the federal advisory committee act and tasked with evaluating the ROP’s pilot program.² UCS has testified several times before Congress about the ROP’s pros and cons. Over the past decade, UCS has continued to monitor the ROP.

The ROP uses seven “cornerstones” to describe the essential features of reactor safety, radiation protection, and security. Performance in these cornerstones is assessed on a quarterly basis using nearly 20 discrete performance indicators reported by the reactor owners supplemented by findings from NRC inspections.

The link between the assessment component of the ROP and mandated NRC responses is called the Action Matrix. The Action Matrix, shown in Table 3 below, features five columns. When the performance indicators and inspection findings all fall in expected ranges, a reactor is placed in the left-most, or “Licensee Response,” column of the Action Matrix, reflecting the fact that the licensee takes responsibility for addressing these minor problems and the NRC continues with its normal inspections. If performance in a cornerstone drops a little below expectations, the reactor moves rightward into the “Regulatory Response” column, reflecting the fact that the NRC now responds by increasing inspections. If performance drops further in a cornerstone or declining performance is detected in another cornerstone, a reactor moves rightward into the “Degraded Cornerstone” column, and the ROP mandates additional NRC inspections. If declining performance deepens and/or broadens, a reactor moves into the “Multiple/Degraded Cornerstone” column and the NRC takes further action. If performance problems reach epidemic proportions, a reactor enters the “Unacceptable Performance” column and is shut down by the NRC.

To lead off with the best news, the 10 reactors listed in Table 1 resided in the Licensee Response column of the Action Matrix for the first 10 years (40 quarters) of the ROP’s existence. It does not take much of a performance drop to move a reactor from the Licensee Response column into the Regulatory Response column, so the owners of these reactors deserve recognition for the sustained effort this outcome reflects.

The top performers were evenly distributed among the NRC’s four regions: regions I and III each had two reactors while NRC Regions II and IV each had three reactors.

Table 1: Action Matrix Stars

- Arkansas Nuclear One Unit 2 (AR)
- Catawba Unit 1 (SC)
- Catawba Unit 2 (SC)
- LaSalle County Unit 1 (IL)
- Millstone Unit 3 (CT)
- Monticello (MN)
- North Anna Unit 1 (VA)
- Pilgrim (MA)
- San Onofre Unit 3 (CA)
- South Texas Project Unit 1 (TX)

On the down side, the 20 reactors listed in Table 2 spent considerable time not in the Licensee Response column of the Action Matrix. Seven reactors operated at least half of the past decade with identified performance shortfalls.

| Table 2: Action Matrix Poorest Performers | |
|--|--|
| Reactor | Percentage of Quarters NOT in the Licensee Response Category |
| Oconee Unit 1 (SC) | 67.5 |
| Cooper (NE) | 65.0 |
| Kewaunee (WI) | 60.0 |
| Point Beach Unit 1 (WI) | 60.0 |
| Perry (OH) | 55.0 |
| Point Beach Unit 2 (WI) | 52.5 |
| Oconee Unit 3 (SC) | 50.0 |
| Fort Calhoun (NE) | 47.5 |
| Browns Ferry Unit 1 (AL) | 46.2 |
| Callaway (MO) | 45.0 |
| Calvert Cliffs Unit 1 (MD) | 45.0 |
| Indian Point Unit 2 (NY) | 45.0 |
| Oyster Creek (NJ) | 45.0 |
| Palo Verde Unit 3 (AZ) | 45.0 |
| Braidwood Unit 1 (IL) | 42.5 |
| Palo Verde Unit 1 (AZ) | 42.5 |
| Palo Verde Unit 2 (AZ) | 42.5 |
| Surry Unit 1 (VA) | 42.5 |
| Surry Unit 2 (VA) | 42.5 |
| Donald C. Cook Unit 2 (MI) | 40.0 |

The chronic poor performance of these 20 reactors is troubling. The predecessor to the ROP was the NRC’s Systematic Assessment of Licensee Performance (SALP) process. The SALP process evaluated performance every 18 to 24 months in a handful of broad categories and reported “problem” plants via a “Watch List.” Among its many shortcomings, the SALP process was ineffective in bringing about the needed improvements when performance declines were identified. For example, the Dresden nuclear plant in Illinois was on the NRC’s Watch List for 8 of the SALP process’ final 10 years. The fact that under ROP nearly 20 percent of current operating reactors spent 40 percent or more of the last 10 years in the NRC’s penalty box suggests that neither SALP nor ROP are effective regulatory oversight programs.

An effective oversight program detects flagging performance as early as possible and induces corrections as quickly as practical in order to maximize time spent in the desired performance range. The ROP needs to do a better job of restoring and maintaining performance levels at reactors so as to avoid reactors loitering in the poorer performance columns of the Action Matrix.

A ROP Success Story: Palo Verde

While all three units at the Palo Verde nuclear plant in Arizona spent more than four years of the past decade outside of the Licensee Response column, this result reflects the ROP at its best. When performance indicators and NRC inspector findings first signaled that performance at this site was declining, the plant’s owner, the Arizona Public Service Company, adamantly refused to see the handwriting on the wall. Instead, it blamed isolated occurrences and statistical anomalies for the apparent performance declines. With the additional oversight resources applied when of the reactors’ performance moved it into the Regulatory Response column, the NRC identified several more “isolated occurrences.” As the reactors’ performance dropped into the Degraded Cornerstone and Multiple/Repetitive Degraded Cornerstone columns, the owner replaced the plant’s senior management. The new management replaced

efforts to defend declining performance levels with efforts to improve performance. The ROP raised caution flags early. The plant’s owner could have heeded those warnings earlier and stemmed the performance drop. When it did not, the ROP raised more and more caution flags until the owner stopped and reversed the declining safety performance.

A ROP Failure: Davis-Besse

If Palo Verde showed the ROP at its best, Davis-Besse revealed ROP at its worst.³ In early March 2002, after ROP had been in place for two years, workers at Davis-Besse discovered a football-sized hole in the reactor vessel head. The reactor remained shut down for two years as the owner replaced the degraded head and corrected numerous other safety problems. The NRC estimated that this near-miss at Davis-Besse came closer to a reactor meltdown than any other US reactor incident since the Three Mile Island accident in 1979.

In April 2002 the NRC conducted a public meeting near Davis-Besse to provide its ROP assessment for the preceding year. Under the ROP, the NRC assessed Davis-Besse to be one of the top performing plants in the entire country. In reality, Davis-Besse was one of the worst performing plants of all time. A reliable safety assessment system must be able to distinguish the best from the worst performers.

Message from the Matrix

| Table 3: Average Number of Reactors in Action Matrix Columns, 2000-2010 | | | | | | |
|---|--------------------------|----------------------------|-----------------------------|--|---------------------------------|-------------------------------------|
| NRC Region | Licensee Response Column | Regulatory Response Column | Degraded Cornerstone Column | Multiple / Repetitive Degraded Cornerstone | Unacceptable Performance Column | Other than Licensee Response Column |
| I | 20 | 5 | 1 | 0 | 0 | 6 |
| II | 26 | 5 | 1 | 0 | 0 | 6 |
| III | 18 | 4 | 1 | 1 | 0 | 6 |
| IV | 16 | 3 | 2 | 0 | 0 | 5 |
| All ⁴ | 80 | 17 | 4 | 2 | 0 | 23 |

Table 3 presents the Action Matrix results from the ROP’s first 40 quarters. The average quarter featured 80 reactors in the Licensee Response column, 17 reactors in the Regulatory Response Column, four reactors in the Degraded Cornerstone column, two reactors in the Multiple/Repetitive Degraded Cornerstone column and no reactors in the Unacceptable Performance column.

Thus, each quarter averaged 80 reactors being in the Licensee Response column and 23 reactors not being in that column. NRC Region IV averaged five reactors while NRC Regions I, II, and III averaged six reactors not in the Licensee Response column per quarter.

This regional uniformity has good aspects as well as bad aspects. On the plus side, it strongly contradicts the notion that NRC really stands for “Nielsen Ratings Commission” with the agency’s actions (and inactions) being dictated more by media than merit. The reactors in NRC Region I (notably Indian Point, Oyster Creek, and Vermont Yankee) receive a disproportionate amount of media coverage and Congressional interest. But despite the large spotlight on Region I reactors and smaller penlights on Region II, III, and IV reactors, the Action Matrix results for all regions are roughly the same. These results do not support the belief that NRC is playing to the crowd.

While it's good that media and Congressional attention results do not appear to unduly influence the process, it's bad that the uniform distribution may result from other undue influences. Circumstantial and anecdotal evidence suggest that NRC's inspection resources may be playing too large a role in ROP outcomes.

Table 3 showed that the average quarter had 23 reactors not in the Licensee Response column of the Action Matrix and this result was shared uniformly by the NRC's four regions. The data suggest the ROP may be designed to handle roughly six reactors per region not in the Licensee Response column (a.k.a. "problem plants")—with variation around this number depending on how serious problems at the specific plants are and how many inspector posts are vacant.

This suggests that limits on NRC inspection resources may play a role in deciding when plants are moved to and from the Licensee Response column as part of the ROP process. When new problem plants emerge with problems that cannot be ignored, previously trouble-plagued plants may be suddenly cured and removed from the list of plants that require heightened NRC involvement. And when trouble-plagued plants improve and move into the Licensee Response column, the freed-up NRC resources may allow the NRC to turn to other plants that were previously not on the list but whose problems the NRC can now address. In this way, a relatively equal distribution of NRC resources between regions can lead to a relatively equal distribution of plants that are not in the Licensee Response category.

Inspection resource issues have been previously identified as affecting NRC's oversight decisions and actions. For example, the reasons that the ROP did not detect numerous warnings signs about the significant safety problems revealed at Davis-Besse in spring 2002 included NRC inspection resource limitations. The NRC's own Davis-Besse lessons-learned task force stated in its September 2002 report:

Regional staffing and resource issues challenged the NRC's ability to provide effective regulatory oversight of DBNPS [Davis-Besse Nuclear Power Station]. Because DBNPS was viewed as a good performer, other Region III plants with recognized areas of concern received priority treatment with respect to resource allocation. As a result, NRC inspection and oversight resources provided to DBNPS were minimal during much of the period in which RCS leakage [reactor coolant system leakage—the water leaking through the cracked nozzles in the reactor head] in containment, including its symptoms, were occurring. During the period that RCS leakage symptoms and indications were becoming evident, there were unfilled vacancies for resident and region-based inspector positions for significant periods of time. The regional branch that had oversight responsibility for DBNPS also had oversight responsibility for another plant that was in an extended shutdown for safety performance issues. In 1999, significantly fewer inspection hours were expended at DBNPS relative to other single unit plants within the region. Regional senior manager site visits to DBNPS were relatively infrequent. Two consecutive resident inspector vacancies were filled by NRC staff members who had not previously certified as a reactor operations inspector.

Because the NRC viewed Davis-Besse as a good performer, it reallocated inspection resources from that site to probe bad performers in the region. The drastically curtailed inspection effort at Davis-Besse produced little to no findings, falsely sustaining the notion that it was a good performer and justifying further reductions in the inspection resources.

Not only did NRC's perception that Davis-Besse was a good performer factor into how many inspectors were sent to the plant, but that perception also factored into the grading system. For example, on March 27, 2001, the NRC issued a report on its inspection of the problem-identification and resolution programs at Davis-Besse. Federal regulations require plant owners to implement these programs to find and fix safety problems in a timely and effective manner. The NRC reported:

The team concluded that problems were properly identified, evaluated, and resolved within the problem identification and resolution programs.

The NRC’s report noted that several of the corrective action reports reviewed by the team had to be revised or supplemented to make them complete and accurate, but overall the NRC essentially gave the programs ‘A’ grades. Davis-Besse had to receive good grades—the region lacked the inspection resources to handle another problem plant.

After the football-sized hole in the reactor vessel head was discovered in March 2002, the NRC revisited the grades it had given not long before. The problem-identification and resolution programs now received ‘F’ grades. On August 16, 2002, the NRC notified FirstEnergy, the owner of Davis-Besse, that the reactor would not be permitted to restart until the inadequate problem-identification and resolution programs were corrected. In reality, the problem-identification and resolution programs did not fall apart in a year. They had been flawed for a while, since the problems leading to the hole in the vessel head had existed and gone undetected for quite a while. But NRC’s reduced testing regime and its grade inflation program misrepresented that reality.

Under the ROP, the NRC allowed its perception that Davis-Besse was a good performer to (a) significantly decrease how often it examined conditions at the plant, and (b) significantly downgrade the importance of problems its inspections detected. Davis-Besse was in the Licensee Response column of the Action Matrix because that’s where the NRC believed it should be. Perception rather than performance dictated Davis-Besse’s placement in the Action Matrix.

Figures 1 and 2 show how easily the NRC findings can manipulate the Action Matrix rating of plants, and therefore manipulate the problem-plant inventory. Figure 1 shows the ROP performance indicators for 16 reactors in the 3rd quarter of 2010 (these are just the first 16 reactors in an alphabetical list).⁵ Virtually all the performance indicators (which are provided by the plants) are Green for all the reactors all the time. Green indicates the Licensee Response category, while White, Yellow, or Red indicates a different category. An investigation in 2006 by the Government Accountability Office found that the performance indicators are Green more than 99 percent of the time.⁶ Clearly, the performance indicators reported by the plants are not driving reactors out of the Licensee Response column.

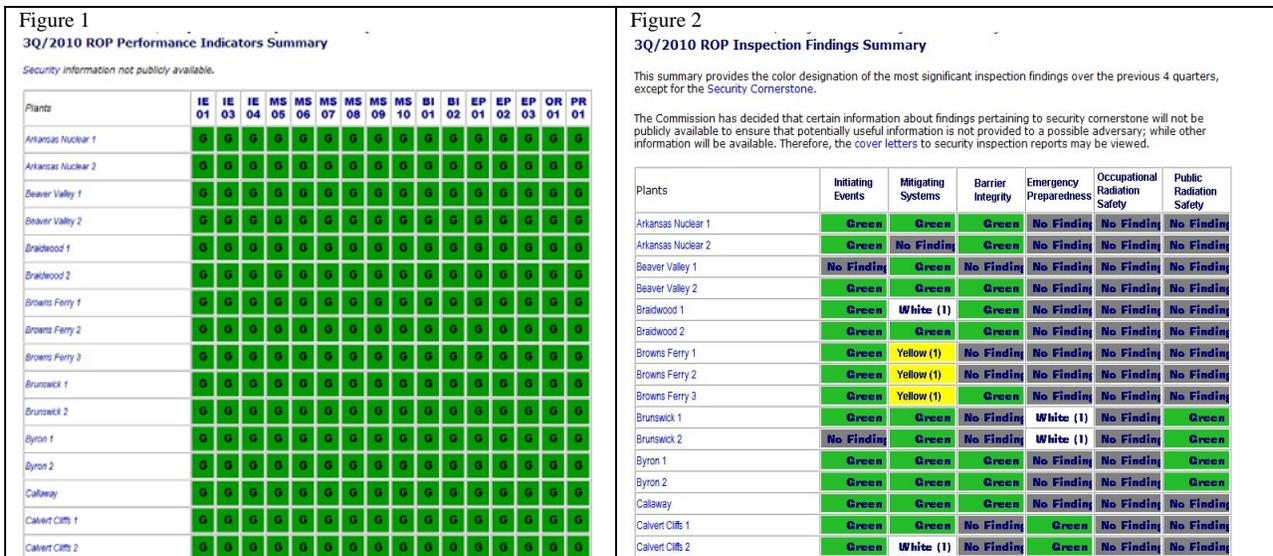


Figure 2 shows the NRC’s inspection findings for these same 16 reactors in the 3rd quarter of 2010.⁷ The NRC generates the findings and determines both their colorization and the timing of that announcement. The NRC conducts the follow-up inspections on an arbitrary schedule to check to see if White, Yellow,

and Red findings have been properly remedied and can change to Green or No Finding status. Thus, the NRC controls when reactors leave the Licensee Response column and when they return to it.

How can NRC's inspection resources dictate reactor placements within the Action Matrix columns? To free up resources when a new inspection finding is about to push a reactor out of the Licensee Response column, the NRC can send a team out to re-inspect findings at a reactor already outside the Licensee Response column. When that inspection gives the reactor a clean bill of health, that reactor returns to the Licensee Response column, ending the need for expanded inspection effort there.

To prevent resources from being stretched too thin, the NRC can either downplay problems found by inspections (e.g., an inspection finding that would be White during a resource-rich period magically becomes a Green finding during a resource-limited time) or delay issuing their findings about problems they discovered for a quarter or two until after problems at other reactors have been resolved, which would free up inspectors. It would be easy for NRC resource issues to play a significant role in determining how many reactors fall outside the Licensee Response column of the Action Matrix.

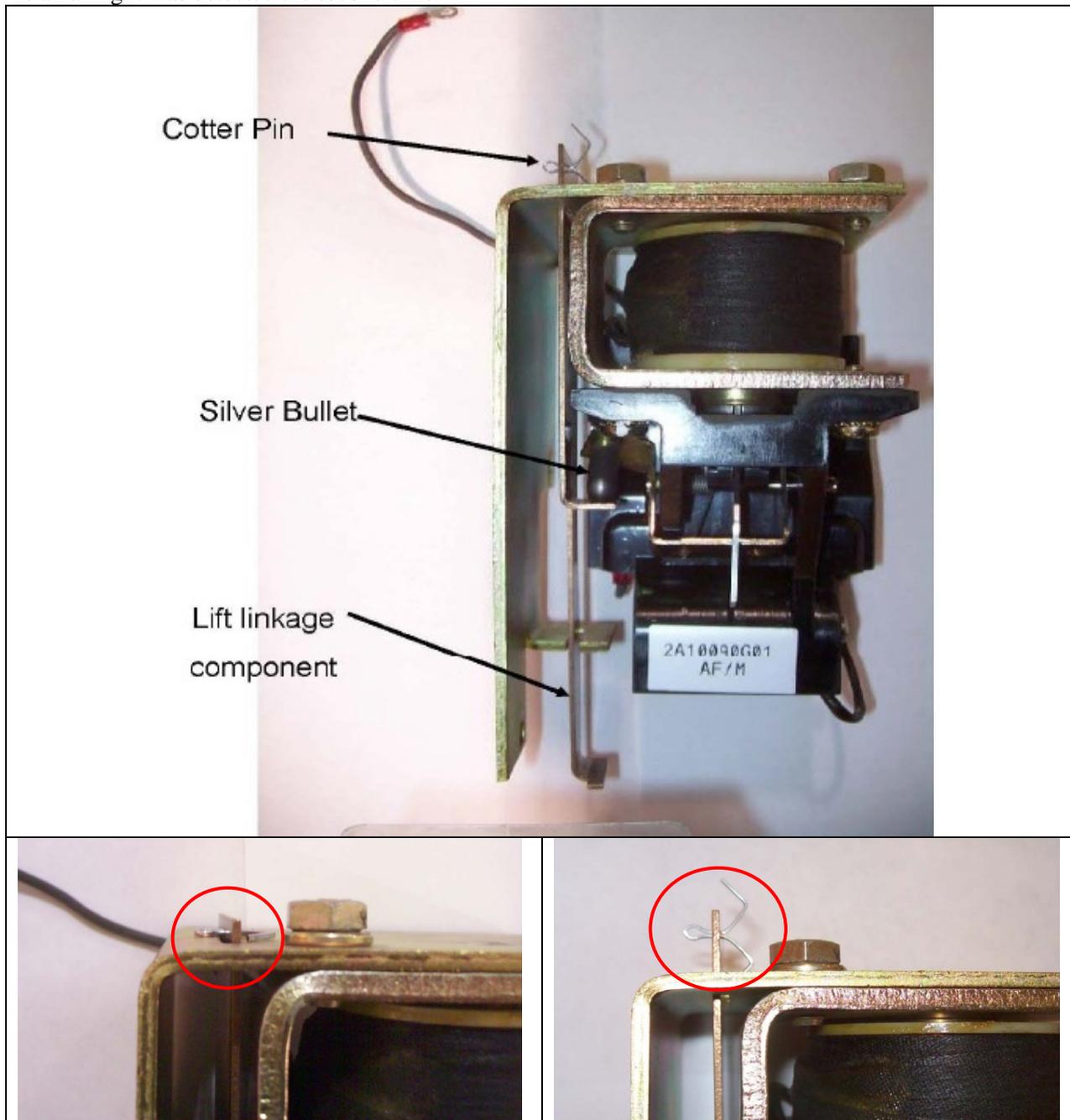
The colorization of NRC inspection findings is subjective. The Significance Determination Process (SDP) is used to evaluate the safety significance of findings and their associated colors. But there are many subjective terms within the SDP that can readily yield any desired color. For example, there are value judgments about whether maintenance workers should have diagnosed flawed components during prior failures and whether operators would have been able to implement alternate measures to ensure safety despite component failures. When an initial SDP fails to match the NRC's preconceived notion of what the color should be for a specific plant, it's a simple task to go back and tweak a term or two in order to obtain the desired color.

Over the years, NRC insiders have told me about instances when NRC management based its decisions about the colorization of NRC findings on resource issues rather than safety significance. For example, NRC inspectors discovered that workers at a midwest nuclear plant had performed inadequate maintenance on a vital safety component, resulting in its repeated failures over roughly a one-year period. The NRC inspectors documented multiple opportunities when workers could have detected and corrected their flawed techniques. Their misses caused this vital safety component to be in service, but unreliable, for nearly a year. The important role of this component in preventing reactor core damage coupled with the long duration that it was impaired produced a preliminary determination that the finding should be Yellow. But a Yellow finding would have driven this reactor out of the Licensee Response Column and required additional NRC inspections. At this time, the NRC's inspector warehouse was empty—inspections at other problem plants had emptied it. So NRC management over-rode the recommendations of its staff that the finding be issued as Yellow and instead issued a Green finding.

I have also heard accounts of findings being changed to reflect more problems when the NRC inspector warehouse was full. In one case, a preliminary Green finding was issued as a White finding because NRC management felt there was something wrong at a reactor and needed more boots on the ground to prove it. The ROP has a process to allow NRC management to do more, or less, than the specified level of inspection effort. Such deviations must be justified in writing and approved by NRC senior management. It is totally improper to bypass this formal process and fudge inspection findings to reach the desired level of inspection effort.

And the NRC issued a White finding late last year for an emergency diesel generator problem at the HB Robinson reactor. An electrical relay (i.e., an on/off switch) problem caused the emergency diesel generator to fail and then fail again months later. The NRC based its White finding on its evaluation that maintenance workers should have figured out the cause of the problem after the first failure and taken steps to prevent the second one. That would have been nice, but the reason for the failure was so subtle

and so nearly random that it is unreasonable to presume it should have been found the first time. In fact, it is amazing it was detected the second time.



The relay that failed is shown in the three pictures. A cotter pin was used to hold a vertical metal linkage bar in place as the relay was being assembled. Once the relay was assembled, this linkage bar was held in place by connections to other parts of the relay. The cotter pin was left installed even though it served no purpose during relay operation.

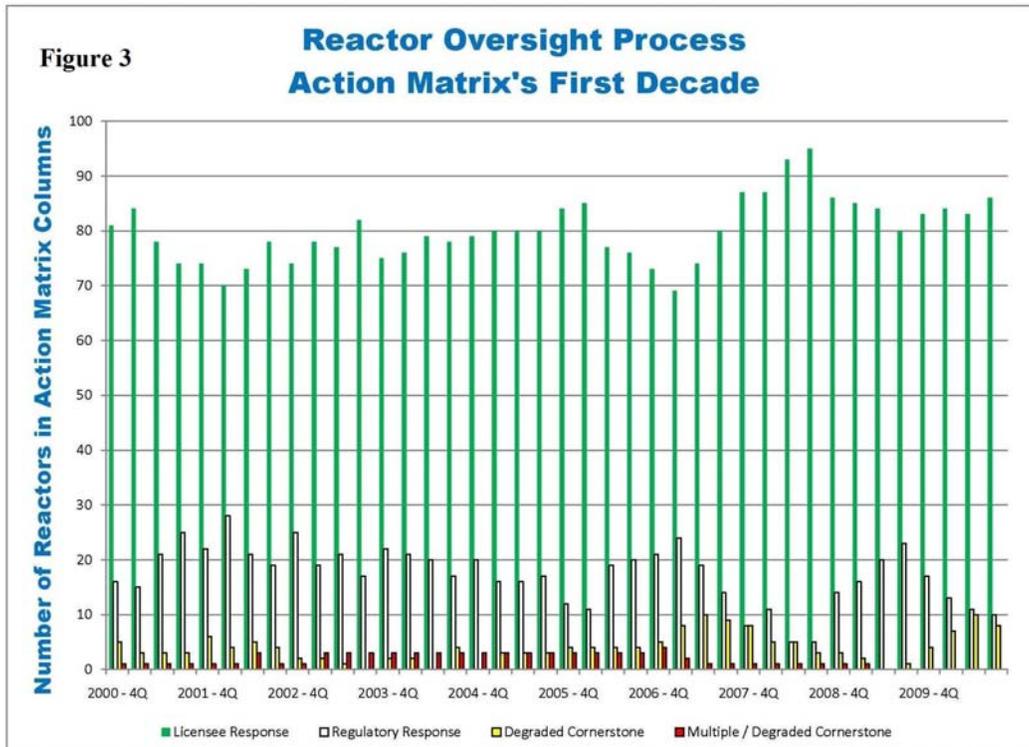
After the second failure, workers found that the cotter pin, when aligned in a vertical position (as shown in the bottom right and top pictures) prevented the linkage bar from fully traveling downward. This travel restriction prevented the relay from functioning properly and in turn prevented the emergency diesel generator from operating.

During the repair efforts following the initial failure, workers operated the relay numerous times and the cotter pin always ended up as it appears in the bottom left picture. The workers did not realize that the cotter pin—about as robust as a thin paper clip or bobby pin—could rotate and, once rotated, remain in that vertical position against the weight of the metal bar. They concluded the relay was not the problem and placed the emergency diesel generator back in service. Even if the workers replaced this relay with a brand new relay, it would not have solved the problem. Only after workers discovered this failure mode following the second event did the vendor re-design the relay to eliminate the cotter pin problem.

The problem itself would likely have led to a “Non-Cited Violation” rating at most had it occurred elsewhere. But the Robinson plant had several problems earlier in the year that apparently caused it to use up its “Get Out of Jail Free” cards. The NRC dialed its color wheel from Non-Cited Violation past Green to White even though this situation did not seem to merit such a finding.

The ROP appears objective on its surface because numbers control it: two White findings trigger one mandated NRC response while one Yellow finding triggers a different NRC-mandated response. Because the performance indicators supplied by the reactors are Green virtually all the time, NRC findings almost exclusively determine the nature and number of ROP colors. The NRC determines—with considerable subjectivity—which colors are assigned to findings, when these findings are issued, and when these findings are closed out.

Figure 3 depicts the quarterly data that produced the results summarized in Table 3. The Action Matrix column inventories are shown for the ROP’s first 40 quarters. Fluctuations in the number of reactors in the Licensee Response column (i.e., the green bars) are seen from quarter to quarter. A reactor in the Licensee Response column receives only minimal NRC inspection effort. That inspection effort steps up significantly as reactors move into the Regulatory Response, Degraded Cornerstone, and Multiple/Degraded Cornerstone columns. When the number of reactors in the Multiple/Degraded Cornerstone column dropped, the number of reactors in the less resource-intensive Degraded Cornerstone column increased by more than the number of those reactors (compare the average heights of the yellow and red bars in the periods from 2003 to 2006, 2007 to 2008, and 2009 to 2010). Coincidence?



Perhaps there are other explanations for some of the events described here. But the possibility that ROP findings are being influenced by resource issues or other factors can't be dismissed out of hand. The NRC headquarter's staff maintains the ROP tools that are implemented by the NRC's regional offices. The NRC headquarter staff should examine the timing of inspection findings that drive reactors out of and back into the Licensee Response column to see if inspection resource issues, or other issues, are having unintended consequences.

CONCLUSIONS

The ROP's first decade had more good than bad. Obviously, that Davis-Besse happened on the ROP's watch is bad. But the ROP functioned well at Palo Verde.

The ROP needs to do a better job of preventing performance-challenged plants from languishing in that condition. It is simply unacceptable that a reactor operate for 67 percent of the decade with identified performance deficiencies. One of NRC's key safety regulations (Appendix B to 10 CFR Part 50) requires that plant owners fix safety problems in a timely fashion and in an effective manner to prevent recurrence. For the same reasons the agency developed this regulation, the NRC needs to resolve known safety performance problems in a more timely and effective way.

Evidence, albeit circumstantial and anecdotal, suggests that the placement of reactors in the ROP's Action Matrix may be influenced—perhaps strongly—by factors other than underlying performance levels.

The NRC should investigate whether resource limitations are affecting what rating reactors are given, and whether there is evidence that those ratings result from considerations other than expert judgment based on the results of inspections.

To ensure that the ROP process works as well as possible the NRC should develop a clear set of meaningful performance indicators against which reactors are evaluated, and increase transparency of the process and the results of the evaluations. The number of reactors outside the Licensee Response column should not be governed by NRC inspection resources. As objectively and transparently as possible, performance as measured by meaningful performance indicators and unbiased NRC inspection findings should determine where reactors reside in the ROP's Action Matrix.

The NRC should develop criteria that allow it to make decisions about allocating its resources and prioritizing which reactors to focus on if the number of problem reactors exceeds the resources it has to apply to them. If the NRC repeatedly finds that the number of problem reactors exceeds its resources, then it should use that to argue for additional resources.

The ROP's inspections are based on a three-year cycle. Some inspections are performed at every reactor every year. Other inspections are conducted at a reactor every third year. The NRC should annually review inspection findings over a rolling three-year period to assess whether the colorizations for similar violations are consistent. For example, that review should compile all the findings from the triennial fire protection inspections performed at all plants during the past three years and evaluate whether the resulting severity levels were commensurate with the underlying safety problems.

Prepared by:
David Lochbaum, Director–Nuclear Safety Project

¹ For more information on the ROP, see <http://www.nrc.gov/NRR/OVERSIGHT/ASSESS/index.html#section1>

² For the panel's final report, see <http://www.nrc.gov/NRR/OVERSIGHT/ROP/ppepfinalreport.pdf>

³ For more information about the Davis-Besse debacle, see <http://www.nrc.gov/reactors/operating/ops-experience/vessel-head-degradation.html>

⁴ Totals affected by round-off to nearest whole numbers. Note that Browns Ferry Unit 1 only operated within the ROP for 13 quarters and that Davis-Besse only operated within the ROP for 27 quarters.

⁵ For the complete performance indicator summary, see http://www.nrc.gov/NRR/OVERSIGHT/ASSESS/pi_summary.html

⁶ See <http://www.gao.gov/new.items/d061029.pdf>

⁷ For the complete NRC inspection finding summary, see http://www.nrc.gov/NRR/OVERSIGHT/ASSESS/pim_summary.html