Official Transcript of Proceedings

NUCLEAR REGULATORY COMMISSION

Title: Advisory Committee on Reactor Safeguards
Subcommittees on Plant Operations and
Reliability & Probabilistic Risk Assessment

Docket Number: (not applicable)

Location: Rockville, Maryland

Date: Tuesday, January 21, 2003

Work Order No.: NRC-742 Pages 1-292

NEAL R. GROSS AND CO., INC.
Court Reporters and Transcribers
1323 Rhode Island Avenue, N.W.
Washington, D.C. 20005
(202) 234-4433

UNITED STATES OF AMERICA

NUCLEAR REGULATORY COMMISSION

+ + + + +

JOINT MEETING

ADVISORY COMMITTEE ON REACTOR SAFEGUARDS

SUBCOMMITTEE ON PLANT OPERATIONS

AND

SUBCOMMITTEE ON RELIABILITY AND PROBABILISTIC

RISK ASSESSMENT

+ + + + +

TUESDAY,

JANUARY 21, 2003

+ + + + +

ROCKVILLE, MARYLAND

+ + + + +

The Subcommittee met at the Nuclear Regulatory Commission, Two White Flint North, Room T2B3, 11545 Rockville Pike, at 8:30 a.m., John Sieber, Acting Chairman, presiding.

PRESENT:

JOHN D. SIEBER	Co-Chairman
GEORGE APOSTOLAKIS	Co-Chairman
MARIO V. BONACA	Member
F. PETER FORD	Member
THOMAS S. KRESS	Member

PRESENT: (CONT.)

GRAHAM M. LEITCH	Member
STEPHEN L. ROSEN	Member
WILLIAM J. SHACK	Member
GRAHAM M. WALLIS	Member

ACRS STAFF PRESENT:

MAGGALEAN W. WESTON

ALSO PRESENT:

CYNTHIA A. CARPENTER

DOUGLAS COE

RONALD FRAHM

TIM FRYE

DONALD HICKMAN

STEVE KLEMENTOWICZ

ROGER PEDERSEN

MARK A. SATORIUS

RANDY SULLIVAN

CONTENTS

P R O C E E D I N G S

(8:32 a.m.)

CO-CHAIRMAN SIEBER: Good morning. The meeting will now come to order. This is a meeting of the ACRS Subcommittees on Plant Operation, and I am John Sieber, Chairman of the Plant Operation Subcommittee, and of the Reliability and PRA Subcommittee, of which George Apostolakis is Chairman. Other members present today are Mario Bonaca, Peter Ford, Thomas Kress, Graham Leitch, Steven Rosen, and Bill Shack.

The purpose of this meeting is to discuss the reactor oversight process as it relates to the Staff Requirements Memorandum, SRM, which directed that the NRC Staff, with input from the ACRS, resolve the apparent conflicts and discrepancies between aspects of the ROP that are risk-informed; for example, significance determination process, and those that are performance-based; for example, those that are based on the performance indicators. Maggalean Weston is the Cognizant ACRS Staff Engineer for this meeting.

The rules for participation in today's meeting have been announced as part of the notice of this meeting published in the Federal Register on

December 27th, 2002. A transcript of the meeting is being kept and will be made available as stated in the Federal Register notice. It is requested that speakers use one of the microphones available, identify themselves and speak with sufficient clarity and volume so that they may be readily heard. We have received no written comments from members of the public regarding today's meeting.

George, do you have any comments?

CO-CHAIRMAN APOSTOLAKIS: No, thank you.

CO-CHAIRMAN SIEBER: Okay. So now we will then proceed with the meeting, and Ron Frahm of the Staff from NRR may begin.

MR. FRAHM: Thank you, John. Good morning. As John mentioned, I'm Ron Frahm from the Inspection Program Branch within the Office of Nuclear Reactor Regulation. Also, as John said, we're here today to discuss the SRM dated December 20th, 2001, and to go over specific concerns that the ACRS identified during our previous briefing on September 9th.

I hope everybody has a copy of the agenda. And if you notice on the agenda, I'm not here alone today. We have several staff members, cognizant experts in their areas, to join me in my briefing

today. These important members of the ROP team include Don Hickman. He'll discuss the Reactor Safety PIs. Mr. Doug Coe will discuss the Reactor Safety Significance Determination process issues. We also have Roger Pedersen to discuss Occupational Radiation Safety. Steve Kelementowicz to discuss Public Radiation Safety, and Randy Sullivan to discuss Emergency Preparedness issues.

I'd like to point out that in the interest of improving the ROP, we actually have an all day Mitigating Systems Performance Index Pilot Program Workshop going on today, as well, downstairs in the Two White Flint auditorium, and it poses a little bit of a problem for us in balancing staff between this briefing and that meeting. And one of the key players is Don Hickman, who I've convinced to stay with us until 10 or 11 today to support all the PI questions, but after that he'll need to go to support the MSPI Workshop, so if we could focus on the PIs as soon as I'm done with my briefing, that would help.

Going to the first slide, we've identified four specific issues from the September 9th briefing that we'd like to focus our discussion on today. First, we'd like to summarize our approach for addressing the SRM that John quoted regarding risk-

informed and performance-based elements, and I will discuss that first this morning. The second and third issues on this slide were specifically identified during -- I'm sorry, not in the briefing, but in the February letter 2002. The risk-informed performance indicator thresholds for the initiating events and mitigating systems cornerstones will be discussed by Don Hickman during the Reactor Safety PI discussion, and the assessment of concurrence findings issue will be discussed by Doug Coe during the Reactor Safety Significance Determination Process discussion.

You had emphasized on September 9th that you'd like to see actual examples presented to you by the cognizant staff members in these areas of greater-than-green findings, and that's why we've presented the agenda the way we have, to have the right people here to address the questions in their areas, so a significant portion of today's presentation is to discuss these greater-than-green examples and their bases across several cornerstones. And we are prepared to discuss the seven examples that were attached to our December 19th paper, and a few others to help demonstrate the basis for their thresholds and our resultant regulatory response.

CO-CHAIRMAN APOSTOLAKIS: Are we going to

discuss the inspection manual that was sent to us a couple of months ago?

MR. FRAHM: Would that be the draft ROP-basis document?

CO-CHAIRMAN APOSTOLAKIS: Yeah.

MR. FRAHM: We're prepared to discuss it. We weren't specifically going to go through item by item, but as issues come up, we'll --

CO-CHAIRMAN APOSTOLAKIS: Because I have a few questions.

MR. FRAHM: Okay. If you could hold those off, I'd appreciate it.

CO-CHAIRMAN APOSTOLAKIS: Sure.

MR. FRAHM: And I actually do have additional copies of several of the documents that we have sent over. We sent over the draft ROP-basis document, and I believe we handed several of those out again this morning. I don't have any more copies of those, so I hope everybody has one.

Second was the NEI 99-02 Performance Indicator Guidance. I have several additional copies here, as well. And probably most importantly was our letter on December 19th that summarized all the issues, and gave our response to you all in writing. And that's essentially -- the briefing today is pretty

much designed after this paper.

Moving along to the staff approach and plans to address the SRM, I actually have a backup slide in your package that has the direct quote from the SRM in case we need to go back to that during the briefing to clarify our discussions today. And I wanted to point out that we intend to address this SRM in our upcoming Annual ROP Self-Assessment SECY paper that's due to be issued by the end of March.

I'd like to reiterate some of the key discussion points provided in our December 19th response. In the development of and the continued refinement of the ROP, we've used performance-based thresholds based on appropriate regulatory response, and we've incorporated risk insights to the extent they were available and applicable. The ROP regulatory framework includes seven cornerstones of safety, and our regulatory response is based on the action matrix with equal weighting to PIs and inspection findings across all seven of these cornerstones. In other words, we treat a white as a white, and yellow as a yellow, regardless of which cornerstone those issues came out of, and whether they were PIs or inspection findings. We perform assessment reviews on a continuous quarterly and

annual basis for all plants and regulatory actions are taken on performance deficiencies as they are identified.

We recognize from the start that these thresholds would likely need to be adjusted as we learn lessons after some run time of the ROP. We continue to adjust these PI and SDP thresholds to ensure a consistent regulatory response, and several of the examples we're going to discuss today demonstrate that.

We also face the continuous challenge to assure that the ROP meets the competing objectives of remaining predictable, understandable, risk-informed and objective in meeting the four strategic performance goals of maintaining safety, increasing public confidence, increasing efficiency and effectiveness, and reducing unnecessary regulatory burden.

MEMBER FORD: Your continuing adjustment.

MR. FRAHM: Right.

MEMBER FORD: Do you review these PIs on a regular basis like quarterly or yearly, and then see if they need changing?

MR. FRAHM: We essentially review the program continuously, and we do an annual wrap-up of

Lessons Learned throughout the year, so we do an in-depth review every year, and we publish an annual SECY paper.

MR. FRAHM: Okay. How is it decided when you look at these whether there is a consistent regulatory response? I'm not quite sure what you mean by "consistent" here. Is it consistency between the ROP and the SDP, or is it consistency among the various colors? I'm not sure what --

MR. FRAHM: It's both.

MEMBER FORD: It's both those things.

MR. FRAHM: The goal being, when you get to the action matrix, you want to treat a white as a white, and a yellow as a yellow. They're all treated equally regardless of where it's coming from, so that's the balance we're trying to maintain.

MEMBER FORD: And how is it you decide whether they're inconsistent or not? Do you have some criteria?

MR. FRAHM: I don't know that we have any specific criteria, but you can identify outliers --

MEMBER FORD: So it's an expert judgment kind of thing.

MR. FRAHM: It's an expert judgment, and there are a few outliers in certain areas, and we'll

actually be discussing some specific outliers we've identified, and what we plan on doing about it later today.

MEMBER FORD: Okay.

MR. COE: I could add to that just a bit. Consistency, another way of talking about consistency of our response is that a 95-001 inspection, which is prompted in the licensee or the regulatory response column of the action matrix is typically between 16 and 40 hours of additional supplemental inspection.

MEMBER FORD: Okay.

MR. COE: Okay. A 95-002, which is prompted by the next column over, is typically between 40 to 240 hours of additional inspection. That's a fairly wide band, but there's that kind of flexibility. And then the 95-003 inspection is typically, in our experience has been anywhere from 1,500 to 2,000 hours of supplemental inspection. That is, of course, the most substantial of the supplemental inspection procedures, so regardless of whether the licensee arrives at that column of the action matrix by either PIs or SDP results, those are the responses that we give, and that's one measure of the consistency that we try to give.

MEMBER FORD: Thank you.

MR. FRAHM: It's a way to focus our resources on the most safety significant issues in the plants with the most significant problems.

MEMBER LEITCH: One thing I noticed in my review of the NRC Web page daily, it turned out that there's announcements of meetings the NRC is going to have with licensees. And on the same day, it just happens, and it just contrasted for me the kind of inconsistent, perceived inconsistency that concerns me. There were two plants, each of whom had two white findings, and the NRC response seemed to be the same. They were setting up to have a meeting with the two plants, and that's what this announcement was about.

One of the plants, I think it was Peach Bottom, the area was emergency planning. And there were two issues there, each of which had generated a white finding. One was an inadequate critique of a drill, and the other was a failure to declare the emergency within the required 15 minutes. Each of those generated a white finding. That was one plant and one reaction.

The other plant I think was Braidwood, and exactly the same reaction, two white findings, same NRC response. But at Braidwood, the problem was an auxiliary feed-water pump that failed to operate under

certain circumstances, and the other was a failure of their corrective action program to properly correct a problem with the safety-related valve. I've forgotten the details of it, but it just seemed to me as I looked at those two cases, and it just happened that they were on the same day so it contrasted them in my mind.

Here we have two plants, each with two white findings, and we're saying, I guess, what - that the safety significance of those things is more or less the same? Because in my mind, it didn't seem that they were.

MR. COE: I would say that what we're trying to say is that we believe that our level of response to those issues should be approximately the same. And we'll have some more examples like that. And then this, of course -- your point is well taken. It's the crux of the discussion that we're having here today. And we hope, at least I hope that success at the end of the day comes from our ability to give you a better understanding of why we think that those kinds of differences, if you will, are still appropriate in terms of how we respond and react. And also, to acknowledge that we don't think that we have a perfect process yet, and we're going to continue to

adjust those thresholds, if we see, or if we believe that, you know, our level is not appropriately matched to the significance of the issue.

MEMBER LEITCH: I'm not saying that the emergency planning issues are not significant, but it seems to me that -- just in thinking about this, it seems to me that the level of significance there is much less than the level of significance with problems with these safety systems.

MR. COE: I understand.

MEMBER LEITCH: Particularly one related to, first of all, a drill critique. In other words, I guess the situation -- and I don't understand all the details, but it seemed to me that they had a drill. The licensee performed a critique. The NRC felt that some issues had been missed in the drill critique that the licensee hadn't picked up, so it seems to me it's an important issue, but it's a level or two removed from the safety system not working properly.

MR. SULLIVAN: I can speak to that, if you'd like to take the time to do that.

MEMBER LEITCH: Yes.

MR. SULLIVAN: I'm Randy Sullivan. I'm the Emergency Preparedness guy, I guess. But the

issue of the critique I suppose is clearer to subject matter experts than it would be to, you know, observers, learned observers. But we changed our process drastically in ROP in emergency preparedness. Perhaps you're aware, but in the previous program, we would make dozens of individual judgments on the performance. We would publish those. We would speak to them in public meetings. It would go in the report. The critique may catch some of them, it may not. We would publish them all.

Under the new program, there's a performance indicator system which captures failures and successes of the most risk-significant areas of EP, and that's the number that you see published, the DEP PI. We backed-off on our inspection. We refocused our inspection program to leave individual performance out of our inspection program. That's now the licensee's purview, and we rely on the PI. And we made some other changes that I won't bore you with. So when we see the licensee miscall a PI hit, they declare a success when it was a failure, it has a greater significance than just missing something in a critique as you're relating. So in other words, it brings into question the efficacy of the PI value.

MEMBER LEITCH: Okay.

MR. SULLIVAN: And that means we have the wrong inspection program. If we can't trust the PI value, then we're doing the wrong inspection, so we ask the licensee to do a root cause analysis to see what went wrong with their critique, so that we can make sure we trust that number, because we look at maybe 10 percent of the opportunities in that PI. Maybe less, it depends on the program. So when we catch a PI being called "wrong", a success when it was a failure, that brings into question the value of it, and hence, we want the root cause analysis on the critique process.

Now is that the same PRA significance as a broken valve that was not found? I mean, maybe not, but the issue is, our inspection program isn't looking at what it should be if we can't trust that number, so it's kind of interlinked.

MEMBER ROSEN: Well, notwithstanding those useful remarks about the emergency preparedness indicator, I think what Graham's point was, was not really answered by Doug. The question that was really posed is, is it the intent of this program to make similar colors mean the same risk-significance, or is it the intent of this program to make similar colors mean the same action by the NRC? And I think it's the

latter --

MR. FRAHM: Clearly, it's the second.

MEMBER ROSEN: -- not the former.

MR. FRAHM: Right.

MEMBER ROSEN: And since it's not the former, any attempts by us to try and rework the ROP to make the colors equal in risk space will be changing the program, since that's not its intent. And that's the difficulty I've had all along with this, that it is true that a white is a white, and a yellow is a yellow, and all colors are equal regardless of which cornerstone they come from, as you said before. But that's only in action matrix space, not in risk space.

MR. FRAHM: Right.

MEMBER ROSEN: And we need to keep that in our minds all the time. And this is the confusion you got into, it sounds to me like.

CO-CHAIRMAN APOSTOLAKIS: But there is a problem with that. The way I understand it, and from Doug's reply and the discussion that followed, the factor that determines, the element that determines equivalence is the response. Okay? We look at two situations and say well, we would respond the same way. We do some investigation that would take about

16 hours or whatever; therefore, they're equivalent. But doesn't that go against the whole idea of risk-informing the regulations?

CO-CHAIRMAN SIEBER: Yes, it does.

CO-CHAIRMAN APOSTOLAKIS: It preserves responses, prior responses and adjusts the colors. Well, the whole idea of risk-informing the regulations is to have a response that is commensurate to the risk level. And I agree with Mr. Rosen, that has been a problem with me from the beginning, trying to understand why these colors are equivalent. And certainly, failure to critique a drill is not of the same safety significance as unavailabilities of safety systems and so on, so we have a fundamental issue here. Are we going to use the response as the criterion of equivalence, in which case, we are really deviating from the idea of risk-informing the regulations, or are we going to use some other criteria like risk to establish equivalence, and then adjust our responses to the risk level?

MEMBER ROSEN: It seems much more intellectually satisfying to me --

CO-CHAIRMAN APOSTOLAKIS: And challenging though. This is really a more challenging --

MEMBER ROSEN: Yes, it is.

MEMBER KRESS: It's extremely challenging because if you're going say from green to white area, it's almost impossible to determine the risk significance of that. Now when you get up to the red area, I'm sure you probably can, but that probably is the only threshold, in my mind, that you can actually establish the risk significance of. So you're stuck with not being able to do what we want to do, and I think you have to then fall back on performance-base in the sense that your thresholds are set by people's judgment. And that's where I think we're having a problem.

MEMBER ROSEN: We live in the real world, and being pragmatic is important, but to -- if we are being pragmatic and not -- and thinking that we're really being risk-informed, I think we're confusing ourselves. And I think it's -- the central element that we're discussing here has confused the ACRS for some time. And I think the staff has been pragmatic about trying to run the ROP in the way they're doing it now, but we need to deal with this from a fundamental point of view.

MEMBER SHACK: Yeah, I mean I have a fundamental disagreement with you. And I don't think that looking at the risk significance is the right way

to look at this. This is a risk-informed process. We are trying to assess licensee performance, you know. That's how we get into this red/yellow threshold at 21 scrams. If you only look at the risk significance of that particular performance indicator, you know, you can run it until hell freezes over. It certainly tells you something about the performance and the attitude of that licensee long before you get to the risk significance. And to me, that's what this program is about, is assessing performance. It's not a safety, you know, a safety status thing. We're not, you know, clicking off, okay, this plant is now at five times ten to the minus four, you know, bing, bing, bing. You want to know something about -- and George, of all people, Mr. Safety Culture Himself, I mean, you know, that's really --

CO-CHAIRMAN APOSTOLAKIS: And a cultured man, of course.

MEMBER SHACK: That's, I think, part of what we're -- you know, we're incorporating things like the EOP. You know, they may not have the same risk significance in the PRA, but they tell you how the licensee's attitudes are, his questioning attitude, his response. There's a lot of these things in the response that I don't think -- you know, that's

my problem with setting -- the risk informed, to me, should be in the selection of parameters. I'm not all sure like Tom, that you can really set the thresholds in a meaningful way by looking at the risk significance of the numbers.

MEMBER KRESS: And I agree with you completely. And I think you have to fall back on just what is our experience, what is our judgment on setting these thresholds. And I think it's a real mistake to mix in in this matrix, here's the performance-based ones, and here's the risk-based ones. I think that's a mistake, and that's where get these big number scrams. We ought to just stick strictly with performance.

CO-CHAIRMAN APOSTOLAKIS: And I fully agree with both of you. I think I mentioned earlier -

- no, I'm serious. If you remember, there were two fundamental problems I had with ROP from the beginning. One was this consistency of colors, and I wrote some comments in the letter. The other one which I proposed here, and of course, it was killed immediately, was that the action matrix mixes indicators that are based on performance with indicators that are based on risk, with indicators

that are based on regulatory requirements. And I propose that we separate -- now that didn't go very far, but I think we're coming back to it now.

I agree that it's a performance issue, so why then should several of these indicators be based on delta CDFs? What kind of performance is that? What does it tell me about performance? Why would the -- you see, on the regulatory limits, maybe there is a point that, you know, if you are above by 25 percent of what the allowed leakage rate is, that tells me something about your performance. But the risk thing with the fundamental program being what we have identified here, that we are changing one parameter at a time, I think we have a problem.

Now my experience in similar issues, you know, in another life, trying to formulate decision-making problems, is that the most difficult part of that is assuring consistency among your attributes. And here, we're just going over it and say well, gee, you know, the regulatory response would be the same so, you know, all whites are the same. Okay? So it seems to me that we have two major problems here, maybe three.

One is, we have to decide what the criteria will be for equivalence, and it could be some

level of performance, deviation from normal performance and so on. And again, as you know, there is the issue of generic versus plant-specific and all that. And second, whether in their action matrix, it's not a completely independent issue, performance and risk should be separate. And the third in my is, you know, Davis-Besse. I'm really disturbed by it.

Now maybe there is another study going on, you know, how the Davis-Besse incident would affect the ROP, but I just don't see how we can claim that this is a successful program when I read in the Chairman's speech somewhere recently that Davis-Besse was green before we found out what was going on. I mean, I just don't see how we can say that. Are we looking at the right things? We really have to put the issues on the table.

And again, I really have to make this clear. I don't want to sound like I'm criticizing the staff. They have done a tremendous job given the pressures they had to produce something, you know, of this magnitude in the time that was given to them. But it seems to me that it's the role of this Advisory Committee to raise these intellectual issues and the foundational issues. It's not our role to ask, you know, detailed questions, although we do that

sometimes too.

MEMBER BONACA: I think one other problem has been for us that in setting the thresholds, an attempt has been made to give it almost a risk-base. You know, it takes that many trips to come to, you know, degradation from ten to the minus five and on, so I think this took us all in the perspective that this was a complete, you know, risk-informed process, and I think only later when we discussed it that we brought up the issue of it is risk-informed in general, but not specifically. It's not risk-based in any way, and really should be a performance process as is. And I think, you know, maybe that's one thing that should be clarified by the staff, to what extent these thresholds have to be, in fact, quantified. You know, that creates a full confusion, I think, by the time, you know, if we commit to doing so. This quantification of how many scrams it takes to degrade from ten to the minus five to ten to the minus four. I mean, when you attempt to do that, you put us on the road to believe that this is a true risk-informed process, and then we try to apply those kind of criteria everywhere else, and we find these disconnects, of course, because you didn't really mean to do it that way. And I think that clarification

would help.

MR. FRAHM: Well, that's really what the first bullet on this slide is getting at, is that the thresholds are performance-based, and we use risk insights to the extent that they're available and applicable, so not every -- for instance, emergency preparedness. There's not a quantitative value you can have for those thresholds. It's strictly performance-based, and based on what we've learned over the years. What makes sense to an expert, to a panel of experts.

MEMBER ROSEN: It could be quantitative. If all had Level 3 PRAs, could we then not quantify even the EP?

CO-CHAIRMAN APOSTOLAKIS: Well, a critique of the drill I don't know.

CO-CHAIRMAN SIEBER: First of all, I agree wholeheartedly with the way that Steve described what the issue is. On the other hand, there are other factors that I think come in, you know, when you talk about emergency planning. A lot of that comes -- is a political issue. It comes from local jurisdictions, the states and public confidence. If public confidence says I want out of here, I want you to tell me when we go, and so that becomes -- that gains more

significance in the entire scheme of things, as opposed to some pump or some valve that's malfunctioning. Although both are important, one has more risk significance than the other. And if you cast everything in terms of total risk significance, then I think that -- and try to work it as a mathematician or an engineer would work it, I think that's where you come up with the problem.

On the other hand, when you say I want the colors and the performance indicators, and significance determination to indicate what I would do under these circumstances. I have a licensee who has done this. How do I respond to that? And use that as the basis to set agency action, then I think that you have a process that satisfies agency goals. But when you go back and say that it's risk-based, you can't. And there we have Bill Shack's argument, there is elements of risk information that are factored in. On the other hand, this is not a risk-based process, in my opinion.

CO-CHAIRMAN APOSTOLAKIS: But what you just said I think is not so consistent. You said you are using the action of the agency to determine, you know, what the color should be. And then later on you said, now I will use the ROP to determine my action.

I mean, that's a little bit inconsistent.

CO-CHAIRMAN SIEBER: Well, but it makes consistency from time one to time infinity

CO-CHAIRMAN SIEBER:

CO-CHAIRMAN APOSTOLAKIS: Yeah, but the --

CO-CHAIRMAN SIEBER: And that's what the process is all about. You know, you wouldn't need an ROP if you had a licensee and only one person committed --

MEMBER KRESS: The trouble is that is the thresholds can converge on just about any number.

CO-CHAIRMAN SIEBER: That's right.

MEMBER KRESS: I mean, you don't have a way for it to converge on what you think is the right number.

CO-CHAIRMAN APOSTOLAKIS: I think we all agree, I think, that the thresholds cannot be risk-based. And that the philosophy here is to look for performance issues.

MEMBER SHACK: Actually, I think the challenge -- the performance indicators, it seems to me, aren't as much of a problem. You know, we can argue over the yellow/red thresholds, you know. Those, to me, aren't even a practical problem. You know, you're not going to get there. The one I have

the more difficulty with is the inspection process, where you focus everything on the SDP, which is risk-based. And I have a harder time coming up with an alternative way to evaluate, and yet, I don't particularly like the answer that I get to, that I look at each individual element and look at its significance which, you know, seems to me have all the intellectual problems I have when I look at a scram system and I say okay, you can scram until this particular indicator gets me into deep doo-doo in my -

- you know, and I don't like that. Yet, when I get to the inspection process, I don't have a good alternative measure of the significance.

MEMBER KRESS: I think one thing that would help along that line is to quite looking at each of these things as individual elements and think of them as a whole bunch of things that together make up the performance.

CO-CHAIRMAN APOSTOLAKIS: I thought they were doing that.

MR. COE: That's what the action matrix purports to do.

CO-CHAIRMAN APOSTOLAKIS: They are doing that now.

MEMBER KRESS: Yeah, but what you do though, is you go --

MEMBER SHACK: It integrates that at a very high level.

MEMBER KRESS: Yeah.

MEMBER SHACK: You've screened out so much before you get there.

MEMBER KRESS: You've screened out a lot before you get there. And the other thing you do is, you go in and you try to decide whether these are common cause findings or not. And that's where I think you're going wrong. That's a lot of judgment involved there, and I think you should automatically almost assume they're common cause, and just treat them all as set things that you look at. And I think that might help. It doesn't solve the whole problem, but that would help.

CO-CHAIRMAN APOSTOLAKIS: Well, what's wrong with having a two-pronged approach? One would be based on performance as it is defined by the PIs, and another one will be a natural extension of the accident sequence precursor program to lower levels of risk. The ASP now looks at significant events, and publishes, you know, events that go to core damage frequency of ten to minus three or something

thereabout. What this is doing now is extending that to lower levels, and says what we found in this plant creates a delta CDF of ten to the minus five or four, and we may want to do something about it. But let's not mix that with the performance part, which is somewhere else. And I don't see what the compelling reason is for us to have a single action matrix. I just don't see it. And I don't think it's revolution. I think a lot of the work has already been done.

MEMBER ROSEN: And in fact, your point about the workshop that's going on contemporaneously with this meeting; there, the risk-informed and the risk-based parts of this program are moving forward with an improvement, in my view, of the main thought about for the performance indicators. We don't have any similar kind of improvements being thought of that I know of in the performance-based side, so these things seem naturally to be moving on separate tracks that we somehow have glued together. And every time we have a problem, it's about this gluing process that doesn't seem to work for us. Its artificiality keeps coming through in our reviews.

MR. COE: I'd like to offer just another thought here, because a lot of what we're discussing revolves around a presumption that performance-based

and risk-informed are somehow really separate and distinct. And what we're tried to do, I think, at a high level kind of philosophically is, you know, the PIs, for example, are measures that are countable. I mean, a good performance indicator is something that's relatively objective, and you can count. That's performance, and when it's possible to do so, we try to set the threshold in a manner which reflects our understanding of the potential risk significance, and that's risk-informed.

In the SDP arena, you know, we've got everything that's -- every inspection finding starts with a performance deficiency. That's performance. We make that conclusion that there is a deficient performance aspect that has had some impact on the plant's, you know, ability to function, and to mitigate, so forth. We make that decision right up front, and then we proceed again to risk-inform what the impact has been.

Ultimately, it's all trying to become more predictable and more objective, and that was what we were trying to achieve over and above what we had in the earlier program. And the point that was made earlier about risk-based versus risk-informed is an important one, and it's been the subject of

considerable debate and dialogue within the staff.

The Commission has spoken on that, and has laid out a definition, but it hasn't helped very much, and perhaps it can be improved in the future. But what I would offer is that risk-informed is a spectrum, and I don't think there's a clear dividing line. This is a personal view now, that there is no clear dividing line between what's risk-informed and risk-based. I think there is a spectrum of being risk-informed, and much of that variation in risk-informed depends on how well the decision stakeholders understand the assumptions that are built into that risk evaluation, and to the extent that they can accept those assumptions as being legitimate and adequate representatives of the situation that's in front of them. So, you know, at the extreme you could say that a risk-based outcome is one in which a number is produced, and a number is, therefore, used by the decision makers without further exploration of the assumptions that stand behind that number.

I would say that that sort of is a definition, a working definition that I would use as risk-based. And I submit that that's not our process in any event, that our process is risk-informed, and we can discuss where we are in the spectrum of being

risk-informed, but I would submit that we have a risk-informed and a performance-based process to the extent that we can bring those things together, so I would just offer that as a thought because it gets to some of the points that are being made.

MEMBER SHACK: It seems to me part of what we're trying to do with this process is to pick up what the PRA misses. And the PRA is very good at looking at the effect of the design, and what happens when equipment goes out of order, the effectiveness of procedures. It's not very good in telling you is the organization prone to having latent errors. You know, does it have a questioning attitude when things aren't exactly the way they are, and you're trying to rationalize for why, what happened. And however we risk-inform it or risk-base it, PRA is never going to tell us about those kinds of things, and so focusing our process too much on that I think misses the other part, and that's the part that I'm worried about.

MR. COE: As are we. And because the earlier comment about the Davis-Besse lessons learned, indeed are having an impact, or will have an impact to some extent on the reactor oversight process. We've been given a number of things to think about and look at. And the philosophy, of course, was in order to

become more objective, we look at the things that have actually occurred that we can count, we can measure, we can analyze to some degree, and represent that as some kind of an impact on the public's health and safety risk. And where we can't do that, we establish some commensurate levels of response so that we would react in a way that we think is appropriate, and we acknowledge that there is a difference there. But ultimately, those three crosscutting areas get to the -- one of which is the safety conscious work environment, gets to the point that you're making.

The assumption originally was that if there are problems in that area, they will reveal themselves through things that we can see, and the expectation was that we wouldn't get the most significant thing that we see right away.

Now perhaps if we, and this is speculative, perhaps if we'd had more opportunity under the reactor oversight process with plants like Davis-Besse, we might have started to accumulate some issues that we were beginning to see at the lower levels before we saw the big one. And I guess we can speculate, but that's all it is.

The point is, is that that was an original presumption of the ROP design. It may change over

time as we continue to reflect on the lessons learned from Davis-Besse, and we're doing that.

MEMBER BONACA: To what extent does the inspection process reviews cause the root cause evaluations at different plants?

MR. COE: Well, that is the focus of the supplemental inspections. When you look at the inspection procedures that I referenced earlier, you'll note a strong emphasis on examining the licensee's root cause of failure, and we make a judgment, an assessment of that in those supplemental programs. Since there has been an issue that has risen to some level, some threshold that we believe further involvement on our part is necessary, that involvement goes to the adequacy of the licensee's own corrective action processes.

MEMBER BONACA: Because often times, I mean, you know, if you really go through them and you have a degraded process, you find that there are latent issues built right into the -- for the process which are not identified by an adequate root cause process, so I'm trying to understand how you do that linkage, and how much the NRC is looking into that.

MR. COE: Yes, sir. That is a focus, supplemental inspection.

MR. FRAHM: In addition to the supplemental inspection, we have a corrective action look built into our baseline inspection program, as well, at all sites, and that's continuous. That's built into each inspection procedure, and we also do a periodic in-depth review of every licensee's corrective action program.

CO-CHAIRMAN APOSTOLAKIS: But the problem with root cause analysis is that there isn't really a universally accepted definition of what is a root cause. And, in fact, it would be interesting to go and pick up some of the AIT reports that the staff has prepared after some serious incident, and where the staff identifies problems with a licensee, and see whether earlier root cause analysis mentioned those. For example, if you read the Davis-Besse investigation report, they talk about I think isolation, of the staff of Davis-Besse not appreciating experience in other facilities. I think the questioning attitude is very astute, but I'm not sure.

I just can't imagine that an engineer doing a root cause analysis for a lesser instance would go down to that level, so I don't know how much value these root cause analyses have if we have not identified what the root cause is. Would these go --

I mean, my colleagues here who have actually worked at the facilities, would these analyses go down to organizational issues? Probably not.

MEMBER BONACA: Not necessarily. Sometimes they do, but I think that typically, you know, if you have problems, for example, in maintenance, the way you do things, and they may result in common cause problems because you do the same, you know, kind of maintenance on a reactor coolant pump or some pump, and then you do it on the others, and then you find that you have root cause evaluations that really don't go deep. They'll ask those questions you cannot trace back to the maintenance process what should have been traced at that level. That's really where you begin to see significant problems, and potential cascading effects in common cause, so that's why I was asking --

CO-CHAIRMAN APOSTOLAKIS: Somebody ought to look at it.

MEMBER BONACA: Yeah, because I mean, when you have then a significant problem at the plant and you get on the root cause process, and you begin to investigate, you find superficiality in so many of them. And you're saying how come you didn't ask this question. And, you know, there is people who are

becoming specialists in looking at those root cause evaluations and looking at, you know, this staircase as you call it, the why staircase. Why did you stop here? Why didn't you ask the next question and so on? And I think, you know, maybe looking into that process gives you some insight. You'll know ahead of time what the culture organization is what potential late issues are.

MR. COE: I agree.

MEMBER ROSEN: The disconnect that we have today on the table in front of us is that you said that you did use all our skills in looking at Davis-Besse's corrective action process, and yet presumptively if that had been done, one would say corrective action process at Davis-Besse is not working well. Therefore, we have a problem long before we had the material defects we found on the reactor vessel head. And so that's the part that doesn't work for me, and says yeah, we were looking at Davis-Besse's corrective action process. Well, then it seems like it ought to have found the lack of questioning attitude across the board, and these corrective action documents that weren't acted on, and all the other things that were later, that have become known. So I'm a little troubled by the idea that the

ROP is okay, we don't have to do anything with it because we did look at corrective action at Davis-Besse. Well, if you did, then we got the wrong answer.

MR. FRAHM: I'm sure there will be several lessons learned from the Davis-Besse recommendations from the task force that we'll incorporate over the next year.

MR. COE: That's right. We're not saying the ROP is okay necessarily, that it can't sustain continued evolutionary improvement. That's certainly part of our objective, and we will be looking at how we can improve relative to Davis-Besse. And I think that the corrective action, or I should say the problem identification and reporting inspections that we do at plants can continue to improve, and the manner in which we can seek out and find these more pervasive problems in licensee corrective action programs, I think there's more to do in that area.

CO-CHAIRMAN APOSTOLAKIS: Are we at some point going to address the issue of performance versus risk? I mean, we raised the issue, but I don't hear any response.

MEMBER ROSEN: I think we should in the letter, if we write a letter --

CO-CHAIRMAN APOSTOLAKIS: Well, I'm asking the staff whether they plan to say anything about it.

MR. COE: The point that I made just a moment ago regarding the -- we believe that we have a performance-based and risk-informed program, that there's an appropriate melding of those concepts in our program.

CO-CHAIRMAN APOSTOLAKIS: I see.

MR. COE: Is really our -- trying to help you understand where the philosophy was, where it came from, and how we're applying it. It, of course, is up to you to decide whether or not you'd like to offer your, you know, recommendations to do something different.

MR. FRAHM: And as the day goes on and everybody does their parts of the presentation, I'm hoping that it becomes more clear. And if we need to revisit this later in the day, we could do that, as well.

CO-CHAIRMAN APOSTOLAKIS: Okay. One last question before we -- we're still at the slide. Is it, on the action matrix colors that talk about two whites or a yellow and so on, how often are these used? How often do you find that you have two whites, or is it the overwhelming majority of cases you have

one white, and you take action and that's it?

MR. FRAHM: Well, there's -- I don't have the number off the top of my head, but there's been several instances where we've had multiple whites.

CO-CHAIRMAN APOSTOLAKIS: Multiple whites.

MR. FRAHM: Sure.

CO-CHAIRMAN APOSTOLAKIS: And these were due to the fact that you are carrying over some incident for several quarters, or in the same quarter?

MR. FRAHM: Well, with PIs the results are what they are, they're indicators of performance. And when a PI changes quarterly, it could go on or off the color threshold.

CO-CHAIRMAN APOSTOLAKIS: Right.

MR. FRAHM: But with the significance determination process, once you cross the threshold and get a white issue, for instance, it stays white in the assessment process for at least a year.

CO-CHAIRMAN APOSTOLAKIS: Right.

MR. FRAHM: Up until the corrective actions are satisfactory, and a few other criteria that we go by.

CO-CHAIRMAN APOSTOLAKIS: So you may have two whites because of this fact.

MR. FRAHM: Right.

CO-CHAIRMAN APOSTOLAKIS: I mean, something happened in January, and something else in September. But to get two whites in the same quarter --

MS. CARPENTER: I can --

CO-CHAIRMAN APOSTOLAKIS: Yeah. Go ahead, please.

MS. CARPENTER: Okay. I'm not sure -- this is Cindi Carpenter from the staff. There is a backup slide, Ron, number 32, where the corner I don't know the answer to that, but over the year we know that for six -- for 2002 we know that two plants reached the degraded cornerstone, which would mean two whites in the same cornerstone.

CO-CHAIRMAN APOSTOLAKIS: Did you say slide 33?

MS. CARPENTER: Slide 32, right.

CO-CHAIRMAN APOSTOLAKIS: 32.

MS. CARPENTER: The backup slides.

MR. FRAHM: And we'll get that up on the screen here.

MS. CARPENTER: The regulatory response --

CO-CHAIRMAN APOSTOLAKIS: Oh, you mean now.

MR. FRAHM: Up here. It's in your slide

package. We're trying to find it.

MS. CARPENTER: No, I'm sorry. But what that slide would show is that for those plants that had two whites that co-existed in the same cornerstone, that would put them into the degraded cornerstone. And our slide for last calendar year for ROP 3 is showing six plants reached the degraded cornerstone.

Now there were a number of other plants that reached regulatory response -- at least one white, or maybe two whites in different cornerstones, which would be the 30. But two in the same cornerstone would be six for last year.

CO-CHAIRMAN APOSTOLAKIS: But this is due to the fact that you are carrying over a color for a period of time.

MS. CARPENTER: Right, for four quarters.

CO-CHAIRMAN APOSTOLAKIS: For four quarters, and I was wondering whether you can get two whites or a yellow in one quarter.

MS. CARPENTER: You could.

MR. COE: Yes, they don't have to initiate that same quarter.

MS. CARPENTER: Right.

MR. COE: If you have a white inspection

finding in the third and it is residing in the action matrix for four quarters, on quarter three you might have a PI pop up as another white on that quarter. That plant is in the degraded cornerstone.

MEMBER SHACK: I think what George is looking at is the number of times you actually have to deal with a simultaneous, you know, that quarter --

CO-CHAIRMAN APOSTOLAKIS: Yeah, the third quarter.

MR. FRAHM: Two new issues showing up at the same quarter.

CO-CHAIRMAN APOSTOLAKIS: Yeah.

CO-CHAIRMAN SIEBER: Well, Graham's example was one of that type, two issues in the same cornerstone.

MEMBER LEITCH: I don't have the timing of those yet. There were two white issues, but I don't have --

MR. FRAHM: It certainly could happen, and I'm sure it has happened, but I don't have a specific example.

MS. CARPENTER: Roger has the --

MR. PEDERSEN: Yeah. This is Roger Pedersen of the staff. I'll be talking to you a little later on about the Occupational Radiation

Safety cornerstone. The example that we're using there in ALARA actually was two white findings in the same outage, the same inspection report. And we recently completed enforcement action for Davis-Besse for the Radiation Protection issues at Davis-Besse, and those are going to be two white findings in the same outage, as well, so it does happen.

MEMBER BONACA: I just have a question before you -- we at some point talk about the issue that Dr. Apostolakis brought up at Davis-Besse, I mean, the issue of you do have a cornerstone which is called barrier integrity and, however, it didn't pick up Davis-Besse before or after. The issue that maybe what you have to look at is the inspections and the quality of inspections. I mean, I'm trying -- I'm wrestling with that issue -- for example, I'm wrestling with the issue, should I see the V.C. Summer event where they missed their ISI existence of cracks as a failure of barrier integrity?

MR. COE: Yes.

MEMBER BONACA: Okay. And then how would I skill my inspection process to pick up those kind of indications? And the reason why I'm worrying about that is that Davis-Besse is another example of that in a way, and to what extent does the ROP get involved

into the inspection process? We have shorter and shorter outages. That's going to be probably a place where you are going to have repeat events of this nature. Unless you look into it, you're not going to see it. And I'm trying to understand to what extent the staff is looking at this issue of using the barrier integrity as a means of monitoring these kind of situations.

MR. COE: Well, we do have an inspection procedure that looks at in-service inspection activities that the licensee performs, and much of that inspection is performed during the outages when the information becomes available to us. We sample a number of different packages that the licensee has either done repairs or done testing, ISI testing. And so there is a basic element of our baseline program to look at that.

Now we modified that procedure after Indian Point tube rupture, because it involves, you know, the steam generator tube integrity inspections, as well. And it looks like it'll be a focus of our attention for -- after we reflect on the Davis-Besse lessons learned, so you may see some additional changes to that procedure. But that is an element, an important element of our baseline program.

MR. FRAHM: Okay. If possible, I'd really like to get through these next two slides, and get over to detailed PI discussions, because as I mentioned earlier, Don Hickman needs to leave us in the not too distant future. But actually, a lot of our discussion over the past several minutes has been on this third bullet, where we're competing with certain goals within the ROP. And a good example is if you're trying to get as risk-informed as you can possibly be, you're losing some of the understandability, some of the public confidence because they just, you know -- the people who are deep in the process may understand it, but those who are looking from, you know, just a general public perception standpoint, they might be missing the boat, so it's a very careful balance. And we struggle with each change we make to the process.

And the last bullet just points out that we recognize that the ROP is not a perfect process. We think it's a very good process, and we do continue to make improvements through our self-assessment and feedback processes, and we continue to have interactions with our stakeholders, including the public, our regional offices, advisory committees like yourself, and the industry. And, in fact, we just

completed internal and external surveys of our stakeholders, and we're in the process of reviewing those surveys, and gathering lessons learned. And we plan to address those in our upcoming ROP annual SECY paper.

The next slide. The SRM, as you're well aware, did request that we provide recommendations for resolving the apparent conflicts and discrepancies between aspects of the ROP that are risk-informed and those that are performance-based. And as we've been discussing for quite a bit, those two terms are not mutually exclusive, and we tried to combine them to the extent we can in the process. But in a nutshell, our position is that the ROP is working effectively today, and that in general, plants are receiving the appropriate level of oversight. And we're making the second statement that plants are receiving the appropriate level of oversight based on our last two agency action review meetings. Our senior level managers all got together and reviewed the plants that are in the higher levels of the action matrix, and they all agreed that they were able to focus their resources on the appropriate plants and issues.

And most recently during our mid-cycle reviews, the regional offices gave us the same

feedback, that they are able to focus their resources on the plants that they feel have the most significant problems.

We also recognize that there are acknowledged differences between risk-informed and performance-based aspects of the ROP, but we consider these differences, and not necessarily discrepancies. And all inputs in the assessment process are performance-based, but some are more risk-informed than others based on the availability of the information and the applicability of the risk information. And we believe that the ROP does effectively address both risk-informed and performance-based issues.

We further recognize the need, and we have for quite some time, that we need to consolidate our basis for these SDP and PI thresholds into a single document, and that's the whole gist of that ROP basis document that we provided a draft of a few weeks back. And we really hope that that goes a long way to consolidating our basis in a more transparent manner, and hopefully making the whole process more understandable.

We do expect continued incremental improvements, as I mentioned on the previous slide

actually, and we do anticipate several upcoming changes as a result of the Davis-Besse lessons learned task force, as well as the SDP task group that was looking at some problem areas in the SDP process. But those reports I don't -- I think they're actually both out there, but we have not delved into them, and really addressed the recommendations, but we anticipate significant changes to the process as we go forward.

And lastly, as we mentioned in the paper and during the September 9th briefing, we have begun discussions with the Office of Research to explore the use of formal decision analysis within the Reactor Oversight Process, but this is very much in its infancy, and this would be considered a long-term project. And as I said earlier, we believe the process is working effectively today, but this might be an area that we'd like to explore as potentially adding some more structured theory to the ROP.

And that's really all I had with regard to the SRM. As we go -- as I said earlier, as we talk more today, I'm sure more issues will come up, so please feel free to ask questions as they do come up on the SRM and how we plan to address it. And with that, I'd like to turn it over to Don Hickman to

discuss the Performance Indicators, and the specific issues of the Risk-Informed Performance Indicator thresholds will be one of the main discussion points.

MR. HICKMAN: Thank you, Ron. We've had good discussion I think this morning about performance-based and risk-informed and that sort of thing. This slide is simply to reiterate, I think, what we've all understood from that discussion, that all of the performance indicators are performance-based. We are counting numbers of particular types of events.

What we've tried to do is to risk-inform those indicators where we could do that. And, of course, the areas most susceptible to that are in the initiating events cornerstone, and the mitigating systems cornerstone, so we have done what we could along those lines.

Of course, when we did that, we used some generic plant models, about a dozen of them, and then we applied that across the industry trying to be conservative with the results of those models. And so they're not maybe the right numbers for every plant, but they should be conservative numbers. And we've had a lot of comment in the past about how we should have plant-specific thresholds. And I think we've

acknowledged that we would like to make the PIs as plant-specific as we can, keeping a few principles in mind; and that is, that the PIs do need to be simple. They need to be something that are clear as to what counts and what doesn't count. Some licensees have no questions. We don't get inundated with questions about whether certain events should count.

CO-CHAIRMAN APOSTOLAKIS: Why can't we use the goals that the licensees have set under the maintenance rule as some sort of threshold for maybe the green/white for the ROP? That would make them plant-specific, and it wouldn't cost us anything. WE've done it already.

MR. HICKMAN: What we are doing is rather than requiring licensees to have PRAs, as you all know, we have developed our own models, and that's what we plan to use for that purpose, rather than relying on the licensees models. We've not checked the accuracy of their PRAs. We've not --

CO-CHAIRMAN APOSTOLAKIS: Well, the maintenance rule is a rule. I hope the numbers they're proposing are meaningful. It's not something they are doing in their spare time.

MR. HICKMAN: The maintenance rule, you're right. I mean, it is a rule, and they are verified by

the residents at the site.

CO-CHAIRMAN APOSTOLAKIS: Right. And so why can't they be the green/white thresholds for the mitigating systems?

MR. HICKMAN: As I say, we have been in the process of developing the SPAR models, and that is what we want to use to confirm the accuracy of the licensee is using. I understand what you're saying.

MR. COE: I think the answer to your question is it could be done that way. And, in fact, I will tell you that that discussion occurred in the development and the conception of the ROP. And it was decided for a number of reasons, I guess independence being the principal one, that we would not rely upon the licensee's maintenance rule, the risk model that they use for the maintenance rule to base those thresholds on.

CO-CHAIRMAN APOSTOLAKIS: Are you saying the maintenance rule is no good?

MR. COE: Not at all.

CO-CHAIRMAN APOSTOLAKIS: So here is the agency saying we are not going to rely on something that --

MEMBER SHACK: Again, if I'm looking at performance rather than trying to assess the safety

status of the plant, it seems to me a comparison between plants which is really where the green/white threshold comes out now, is a very reasonable thing to do. You know, I would call all this white/yellow/red threshold risk-misinformed. You know, and even setting the initiating event green/white threshold on a risk-significant basis, I would almost call risk-misinformed because again, I'm going to look at a single isolated parameter, important as it may be, out of context. And again, that's not what I'm trying to do here. I'm trying to get an assessment of --

CO-CHAIRMAN APOSTOLAKIS: You are raising two issues. I think the white/yellow/red I agree with you, but the green/white I disagree. The maintenance rule says Mr. Utility, come back and tell me what the unavailability of this safety should be or the safety train. Now those guys went back and they looked at their PRAs. They looked at other things, past experience, so on, and said here is our goal. If we meet this, we are doing okay. And this is plant-specific. Now why isn't that green? Isn't that green the whole idea of green?

MEMBER SHACK: If I was measuring the safety status of the plant, yes. If I'm measuring the licensee performance, maybe not.

CO-CHAIRMAN APOSTOLAKIS: But what they did to compare it with other people is using again an unavailability. It's not that they used some other measure. It's the same measure they're using, but they're using now 102 units as opposed to the specific plant. The fundamental approach is the same. They're using the same metric.

MR. HICKMAN: George, that point has come up in many of our discussions with industry, particularly -- primarily with regard to the safety system unavailability indicator, and we've gotten -- industry has proposed different positions. They would like the indicator -- they were looking at a relationship between the green/white threshold and the maintenance rule requirement. And the discussion was, should the maintenance rule be lower than the threshold so they could fix the problem before they went white? Should they be the same? Should it above that? There's been a lot of discussion about that, as to actually what you would do with that number, where you would --

CO-CHAIRMAN APOSTOLAKIS: What was the conclusion?

MR. HICKMAN: The conclusion was that we don't really necessarily want to relate the PI

threshold to any particular value of that maintenance rule. I think they would like to have the maintenance rule value be higher than the green/white threshold - I'm sorry - lower than the green/white threshold so that they could fix the problems before they go white. That, I think, is their position the last time I think we spoke about this. But we have also discussed doing just that, setting it to be the same. That's a big issue that has been discussed quite a bit in the MSPI as to whether there should be that relationship. And, in fact, we're not doing that.

MEMBER KRESS: But you would have plant-specific PIs then.

CO-CHAIRMAN APOSTOLAKIS: They don't.

MEMBER KRESS: No, but if you tried to do that they --

CO-CHAIRMAN APOSTOLAKIS: Then they would be.

MEMBER KRESS: Yeah. And I don't think -- I think trying to get into plant-specific PIs is going to give you a real headache.

MR. COE: That's precisely what we're trying to do with the MSPI program.

MR. SATORIUS: Yeah, this is Mark Satorius with the Staff. I just came from the MSPI workshop,

and could possibly add a few insights here; and that is, one of the goals of the MSPI was to get around this very issue you were talking about, Dr. Apostolakis, and that was, that we would have a performance index in this case that would be consistent with the maintenance rule, that licensees would not be forced to take two looks at issues via maintenance rule space and PI space to get them interlocked so they use the same criteria. That was one of the things that we're pilot testing this winter for six months, starting in September, so we're --

CO-CHAIRMAN APOSTOLAKIS: So the jury is still out.

MR. SATORIUS: The jury is still out. That's the right answer, yeah.

CO-CHAIRMAN APOSTOLAKIS: I would -- coming back to Bill's point because I think it's important. It keeps coming up. I would agree with you, Bill, if the ROP used a different method to set the green/white threshold, but they're still using unavailability maintenance --

MEMBER FORD: Yeah. Why shouldn't we use the same metric? If you --

CO-CHAIRMAN APOSTOLAKIS: Well, that's my question, why not use the same -- if you're using the

same metric --

MEMBER KRESS: I'm saying use the same metric. It's a threshold. You could have the same metric with different thresholds.

CO-CHAIRMAN APOSTOLAKIS: I don't see why a plant that is highly redundant and this and that has to have the same threshold as a plant that is not. Why?

MEMBER KRESS: Because we're not measuring risk.

MR. SATORIUS: If I'm trying to maintain safety status that would be true.

MEMBER KRESS: That's right.

MR. SATORIUS: If I'm looking at performance, their attitude towards safety the way they're doing it --

CO-CHAIRMAN APOSTOLAKIS: So in principle you are allowing it then if it's very good to drift up, because it's still below the threshold. Right? One of the very good plants at the low percentile can be allowed to have its unavailability of this system go up, maybe by a factor of five or six, and still be below the threshold and be okay, which brings up the other fundamental issue here. Are we comparing with other plants, or are we -- do want to make sure that

the plant as licensed maintains its status? See, these are deeper issues. Well, South Texas would have a field day. They would have high redundancy. They're one of the more recent plants, and now this ROP comes and says we're going to compare you with some of the oldest plants in the United States, so they say great. Okay. So my -- yeah, we could make a lot of mistakes then, and because we're so low --

MEMBER SHACK: Wait until you get an AP 1000.

MR. HICKMAN: Actually, let me just say something about South Texas. They were very concerned about the SSU indicators because they --

CO-CHAIRMAN APOSTOLAKIS: SSU?

MR. HICKMAN: The Safety System Unavailability indicator --

CO-CHAIRMAN APOSTOLAKIS: Oh.

MR. HICKMAN: -- that they're currently using in the mitigating systems because they do a great deal of preventive maintenance, and they said they were going to be close to the green/white threshold just with preventive maintenance, and that it would take very few failures, unavailability hours to push them over the threshold it turns out, so they were very concerned about going white regularly.

CO-CHAIRMAN APOSTOLAKIS: And this is because their preventive maintenance is so strong?

MR. HICKMAN: Right. And we questioned them on that.

CO-CHAIRMAN APOSTOLAKIS: And this is something that we don't want them to do?

MR. HICKMAN: No, no. We questioned them and said do you think you get the benefit out of doing all that much maintenance, and they said we sure do, and we said fine. It's something they choose to do.

CO-CHAIRMAN APOSTOLAKIS: And then we're going to turn around and punish them for that?

MR. HICKMAN: Well, no. They really haven't gone white. Because as you say, they have redundant systems so their concern was unfounded, but they --

MEMBER ROSEN: Four trains of auxiliary feedwater.

MR. HICKMAN: I'm sorry?

MEMBER ROSEN: South Texas has four trains of auxiliary feedwater.

MR. HICKMAN: They have --

MEMBER ROSEN: They have three motor driven and one auxiliary, and one steam driven. So in terms of redundancy, there's a lot more redundancy,

just as an example, in the auxiliary feedwater system in South Texas compared to other auxiliary feedwater systems.

MR. COE: There are similar examples with other plants, as well, plants that are penalized, if you will, for accruing an acceptable amount of unavailability. And because those thresholds were set generically, because back to your earlier point, that was the best we could do right at the beginning of ROP, knowing that we're going to penalizing some plants like that because we set the threshold for the plants with the least redundancy and it would have the most significance if they accrue that level of unavailability. So that's where they would have set, knowing that that was a starting point, and the evolution since then has been towards exploring ways of making that more plant specific.

CO-CHAIRMAN APOSTOLAKIS: So is it possible then that I can have a plant that is a very good performer, and its unavailability for a year of one system or two systems goes up, but it doesn't really reach the threshold because the threshold is very high. But if I look at its PRA, delta CDF is ten to the minus four, is that possible?

MEMBER ROSEN: A very low CDF with a high

unavailability of some safety trains is possible with plants that have high redundancy like South Texas.

CO-CHAIRMAN APOSTOLAKIS: So don't I have some conflict there now?

MEMBER ROSEN: Yes. And that's why there are two South Texas guys downstairs at the MSPI indicator workshop arguing for an even broader MSPI than is being proposed now, to take into account more of the actual equipment than the plant has actually got in place, rather than this artificiality, which penalizes plants with higher redundancy.

CO-CHAIRMAN APOSTOLAKIS: I'm puzzled a little by the Committee's attitude towards something that I think is obvious, and maybe I'm wrong, but I'd like to understand that a little better. Why philosophically is it meaningful to compare the performance of this plant with the whole fleet, versus saying no, we have licensed you. We have agreed with your design, your tech specs and everything. Now the RO people make sure that you stay within a little band there over what we have licensed. Isn't that the whole idea of having an inspection program? Why do I care what happens in Southern California? My plant is here, and I'm -- you know, I have all these rules. I have my PRA, and what the NRC should be saying is

let's make sure that you don't deviate from what we have licensed too much, because that is acceptable. But now we're saying no, we're not going to look at that. We're going to look at you, how you perform compared to San Onofre. For me that's -- I don't see the logic of it.

MEMBER BONACA: I think you're right.

MR. HICKMAN: Let me explain to you, I guess how we got there. We wanted to start the program, get it in place and make improvements as we progressed through the years. For the mitigating systems cornerstone, the data that were available was from the WANO safety system performance indicator. We have that data, and that's what we used. Although it wasn't an ideal indicator, it served the purpose initially. And what we had then was performance across the industry. We chose for the green/white threshold to use -- to identify outliers from industry performance simply because we could do that. We had the data. We could do that quickly and easily, and we could get something in place.

If you look at what's going on downstairs, the mitigating system performance index, and we don't call it an indicator, it's an index. It gives relative change, but what's going on down there is

that the green/white threshold is set at ten to the minus six for all plants. And then plant models are used to determine when they cross that threshold.

CO-CHAIRMAN APOSTOLAKIS: Yeah, but again let's separate the issues here. It should be separate. One is, what is the right thing to do. And the other one is, we did it a certain way because under the circumstances, years ago, blah, blah, blah. I think they should be separate. And the discussion today is not focused on why you did certain thing. I mean, we are not blaming you for anything. We recognize that you were under tremendous pressure to do something, but still, it seems to me, we have to discuss the fundamental issues of what we're trying to do and so on, not why certain things have been done a certain way. And this is what I'm focusing on.

MR. HICKMAN: Well, I think that's --

CO-CHAIRMAN APOSTOLAKIS: For me, the issue of plant-specific thresholds and so on versus generic is still unresolved. I don't understand why I have to compare my plant with somebody else's on the other side of the country, and why should you care?

MR. HICKMAN: Well, I think that's the way we have --

CO-CHAIRMAN APOSTOLAKIS: I mean, when you

have your inspectors there, does your inspector -- forget about the ROP. Does your inspector go and say oh gee, you know, you violated these but ahh, it doesn't really matter because other plants are doing it too? It doesn't make sense. You have committed to certain things, you better comply. And we are looking at you, not at other plants. And by the way, this is the fundamental idea behind quality control in the industry, that you say look, you're my client. Let's negotiate. You want these kinds of tolerances, then I establish a quality control program to make sure that a year from now I'm still giving the tolerances. I'm not asking myself oh, but this other guy in California is outside, so it's okay for me to be outside too.

MEMBER BONACA: Really that's the way that the regulatory system goes, because typically in the South Texas licensee's plants ultimately to accept the tech specs which are pretty consistent with the industry. It didn't say I have, you know, ten of these redundancies, therefore I can lose five. It said simply that you have certain action statements, you know, for individual trains and so on and so forth, which are probably consistent with the rest of the industry.

MEMBER ROSEN: Well, Mario, what happened is that's exactly what they were in 1988, when the first of the South Texas units went in service, and there was a great penalty to South Texas because of that. And over the years since 1988, the tech specs have been revised to account for South Texas' redundancy.

MEMBER BONACA: Yeah. Okay.

MEMBER ROSEN: Based on risk analysis.

MEMBER BONACA: Yeah.

CO-CHAIRMAN APOSTOLAKIS: Well, maybe we can let you go on for the next two minutes before we interrupt you.

MR. HICKMAN: I think that that's the way we're headed, George. That's what we're doing in the MSPI, but there is a problem with that. I mentioned the ASP Program earlier. The ASP Program counts -- identifies events with delta CDF, delta CCDP greater than ten to the minus six. And we established the green/white threshold at ten to the minus six, and it turns out that we may have problems doing that.

CO-CHAIRMAN APOSTOLAKIS: Delta CCDP?

MR. HICKMAN: Uh-huh.

CO-CHAIRMAN APOSTOLAKIS: Really that low?

MR. HICKMAN: Uh-huh.

CO-CHAIRMAN APOSTOLAKIS: Ten to the minus six.

MR. HICKMAN: Yeah. The ASP Program counts anything above ten to the minus six.

CO-CHAIRMAN APOSTOLAKIS: Well, they look at it, but they don't really publish it.

MR. HICKMAN: Yes they do.

CO-CHAIRMAN APOSTOLAKIS: They publish it.

MR. HICKMAN: Yeah. Yes, they do. Ten to the minus six, ten to the minus fifth, and down. And, of course, the important ones are ten to the minus four, but they do count that. Using ten to the minus six, we're running into problems where a single failure can put a licensee across the green/white threshold. And the reason is primarily for very high safety significant and high reliability systems, such as aux feedwater at the new CE plants that have no feed and bleed capability, so that's causing us some problems. That may scratch the whole deal.

CO-CHAIRMAN APOSTOLAKIS: What you just said now raises a question which is similar to my earlier question regarding the maintenance rule. Since the ASP is already doing it, you know, ten to the minus CDP is pretty low. Why do you need the significance determination process? Why don't you

just take the ASP and run with it, and change it as appropriate? I think they're already doing it.

MR. HICKMAN: The primary difference between the ASP Program and the SDP is one of timeliness. In the ASP Program, they go back to the licensee with their results, get to look at it. He provides any comments of plant systems, or procedures or whatever that they may have missed. And it's revised and sent back again, and it takes -- more than a year behind now.

CO-CHAIRMAN SIEBER: Yeah, it's a year.

CO-CHAIRMAN APOSTOLAKIS: Still they have the tools.

MR. HICKMAN: Right. Shall we go on? We told you last time that we would consider eliminating the yellow/red thresholds for the initiating event PIs. The difficulty -- there's several difficulties with that. One is that without a red threshold we would essentially be sending a message that there is no number scrams that we would consider to be highly risk significant.

CO-CHAIRMAN APOSTOLAKIS: Or you can say that you're sending the message that way before something like that happens you will have acted to make sure you will never get there, so it depends on

how you look at the message.

CO-CHAIRMAN SIEBER: Well, that's outside the program.

MR. HICKMAN: And in fact, we do. And in fact, we do. That's not the white threshold, that's the red threshold. We have the white and the yellow. We have the 95-001, 95-002 inspections, and we fully expect that any licensee that gets into the yellow probably isn't going to go operate much longer due simply to its own management, regardless of what we do.

CO-CHAIRMAN APOSTOLAKIS: You know, one of the considerations in your deliberations should be the reasonableness of this. I mean, you can't just -- I mean the point that Dr. Shack raised, we can't just change one element in the PRA and make it so large that delta CDF becomes ten to the minus four. I mean, you have to question whether that is reasonable too. I appreciate the value of communication with the public, but you can't base something on something that doesn't make sense so, you know, you can maybe change your message that, you know, you never get there.

MR. HICKMAN: We understand that concern. We appreciate that concern. The thing is, in PIs we don't know how to factor in other types of potential

failures. It has to be something that is very clear as to what counts, and the scrams certainly are.

CO-CHAIRMAN APOSTOLAKIS: And why do you need the red?

MR. HICKMAN: I'm sorry?

CO-CHAIRMAN APOSTOLAKIS: You don't need the red. You don't need to have that panel there. You just don't let them get there, period.

MR. HICKMAN: Well, from a public perception standpoint, that would be indicating that there's no number of scrams above six that we would consider to be highly risk significant.

CO-CHAIRMAN APOSTOLAKIS: No, it will mean that they will never get there. You will never let them go. Way before then you will take action.

MR. HICKMAN: At what point? That would be the --

MEMBER ROSEN: At greater than six. You don't have to put 23 underneath there. You put greater than six. And then you put red.

MEMBER SHACK: The ninety-ninth percentile.

MR. HICKMAN: So you're suggesting getting rid of the yellow band I guess then.

MEMBER ROSEN: Probably yes. Yes is the

answer.

CO-CHAIRMAN APOSTOLAKIS: Now one other question. We've had problems in the past with timing. Something was of immediate safety concern or was not of immediate safety concern, and apparently nobody ever did anything about it. When you say considering eliminating, when will we have the answer? When is your consideration going to be complete?

MR. HICKMAN: Well, that's what I was going to tell you today. We don't see how we can eliminate that threshold.

CO-CHAIRMAN APOSTOLAKIS: So you have already considered it and decided against it.

MR. HICKMAN: Well, the question is, if we eliminate the threshold, we have no red band. Whereas, we do with everything else in the initiating event cornerstone and mitigating systems, except the PIs that are not risk-informed.

MEMBER ROSEN: You're erecting a strawman and then knocking it down. You will -- what we're suggesting is you do have a threshold. It's greater than six can be red. It's just having the number 23, 21, whatever is on the table now is ludicrous. So what we're saying is don't make an argument that we can't change it because we wouldn't have a threshold.

Put a threshold in, just don't make it the one you have.

MR. HICKMAN: That's a different argument I hadn't heard before to eliminate yellow and use the red. The thing is that, as you mentioned earlier, George, we have tried to risk-inform the process to the extent that we can. And when you look at the number of uncomplicated scrams that it takes to rise to that level, it really is quite large.

CO-CHAIRMAN APOSTOLAKIS: But this is a performance issue. This has nothing to do with risk. It seems to me this -- I have never heard of any plant getting into a risky situation because the frequency of some initiating event. It's always the combination of little things that are put together, and all of a sudden you have a problem, so the frequency, I'm with Dr. Shack. This should not be risk-based.

MR. HICKMAN: We have what we think is a good balance in that regard. The PIs look at individual events because that's about all we can count in a PI. We can't have all kinds of different combinations that they need to count. And we look at events, individual, singly, but we look at the accumulation of those counts over some period of time.

And when they have too many, we take action.

The SDP looks at an individual event to see what happened in that event, and was that single event risk-significant, so we have that balance in the program. And when we look at what it takes --

CO-CHAIRMAN APOSTOLAKIS: Public confidence is not determined only by the fact that you may have a yellow/red threshold. It's also determined by the quality of your analysis, by the reasonableness of your arguments. And to have a 23 or 25 yellow/red threshold undermines, in my view, public confidence.

MR. SATORIUS: This is Mark Satorius with the staff. I think maybe the best approach here would be this is something that would need to be brokered with industry because it's contained within the NEI guidance. Possibly we could place it on the agenda for our next working group meeting with them. And a solution might just as was suggested by one of the members; and that is, you footnote the fact that, you know, the expectation that the staff will take action prior to the number six, or greater than six scrams such that there is no need for a red/yellow threshold. Maybe it's as easy as that.

CO-CHAIRMAN APOSTOLAKIS: Yeah, it is.

MS. CARPENTER: We have -- this is Cindi

Carpenter from the staff. We have an ROP working meeting this month, and we could put that on the agenda.

CO-CHAIRMAN APOSTOLAKIS: That doesn't sound right to me though. Essentially what you're saying is that you want the permission of the industry.

MR. SATORIUS: I would not characterize it that way. I would --

CO-CHAIRMAN APOSTOLAKIS: I know you wouldn't, but it sounded that way.

MEMBER ROSEN: I think the word "brokered" is a little mis -- unfortunate.

CO-CHAIRMAN APOSTOLAKIS: It seems to me that if there is something logical we should do it. Now if the industry has a comment on points to something that it illogical or maybe not practical then, of course, we should listen. But to say that we will consider it together with the industry doesn't sound good to me.

MR. SATORIUS: I think you're right. I think the word "brokered" was probably not the best word.

CO-CHAIRMAN APOSTOLAKIS: Okay.

MR. SATORIUS: In this instance, I would

just say discuss. We would need to discuss with industry. It is their guidance document that we endorse, so there would need to be some discussion engaged with stakeholders.

MEMBER SHACK: It's a guidance document that meets your inspection program.

CO-CHAIRMAN APOSTOLAKIS: Yes.

MR. SATORIUS: PI Program, yes, sir.

MR. HICKMAN: We had a lengthy discussion prior to the start of the program, even the original pilot program as to who should write the document. And recognizing that there would be many changes coming early in the program that are difficult for the NRC to handle in a timely manner, that NEI would write the document. But they are simply documenting what we agreed to in the medians. However, let me say agreed to. We try to reach agreement. If we cannot and this has happened on several occasions, and we feel it's in the best interest of the program, we will tell them that, and they agree that it's our program.

CO-CHAIRMAN APOSTOLAKIS: Have you endorsed the NEI guidance?

MR. HICKMAN: Yes, we do. Every time there's a new revision we send out a Regulatory Information Summary.

CO-CHAIRMAN SIEBER: I think we should move on.

MEMBER SHACK: We could make it a Frequently Asked Question.

CO-CHAIRMAN SIEBER: Well, I think it's not particularly appropriate to have a bunch of footnotes that modify the basic scheme of the ROP, because now the footnotes become exceptions, and they're more important than the ROP itself. And I think that that's sort of a clumsy way to do it, but I think that we're falling behind, and we ought to move on, if we can, so either speak faster, or cut something out, or we shouldn't ask so many questions.

MR. FRAHM: Actually, we're doing quite well in accordance with the agenda.

CO-CHAIRMAN SIEBER: You're undermining my thought.

MR. FRAHM: We actually had scheduled to go up until 11:00 to discuss the PIs, so if we want to move on, I'm sure Don could be useful downstairs at the MSPI workshop.

CO-CHAIRMAN APOSTOLAKIS: Are you saying we're ahead of schedule?

CO-CHAIRMAN SIEBER: Yes.

MR. FRAHM: Yes, we are.

CO-CHAIRMAN SIEBER: Have we concluded the PI section?

MR. FRAHM: I think we've concluded our prepared remarks. If there's any --

CO-CHAIRMAN SIEBER: Well, are there any additional --

MEMBER BONACA: They knew it would take two hours to cover six slides. That's why --

CO-CHAIRMAN SIEBER: Well, are there any other questions on PIs, because this may be a good time to take a break. If there are no further questions, we thank you for that portion. And I think that we could take about a 20 minute break.

(Off the record 10:02 - 10:23 a.m.)

CO-CHAIRMAN SIEBER: Okay. Let us begin or continue.

MR. COE: Thank you, Mr. Chairman. This portion of the meeting gets into two, I think, distinct -- gets to two distinct questions, the first of which is the treatment of concurrent multiple equipment functional degradation. The second of which is a series of examples which we hope will help illuminate or illustrate the reason that the staff believes that the current thresholds are adequate for the purposes of the ROP, again not without

acknowledging that they can continue to improve as we gain operating experience with the program.

But I would offer to start as we have in our package here with the question of concurrent multiple equipment or functional degradations, and how the ROP was modified after its initial start to accommodate these kinds of issues when they come up. We sent you the guidance directly from our inspection, our SDP procedure. And I would just start by asking if you have any specific questions that you would like to get on the table right away, I do want to make this portion of the discussion --

CO-CHAIRMAN APOSTOLAKIS: This is the draft inspection manual?

MR. COE: No, this is in our letter of December 19th, there is enclosure one that is actually taken directly from our inspection manual, Chapter 0609 that deals with SDP. I just want -- you know, this is a question regarding how we deal with concurrent issues. And if there are any specific questions, I do want to address them. And if you have them to put on the table now, I'd certainly invite that.

CO-CHAIRMAN APOSTOLAKIS: I do agree with the comment Tom Kress made, that the decision, what is

independent and what is not is highly subjective, and probably you have to always assume independence and calculate delta CDF given that you found, you know, a set of things rather than splitting them up. I think that is a realistically conservative way to approach it, conservatively realistic. Because, you know, it comes back to the root cause analysis. What is a root cause is not well defined. In one analysis we find them independent, in another we find them deeply dependent.

MR. COE: And I would agree with that. There's clearly some room for judgment here, and all I would offer is that the ROP objectives are met when our judgments are scrutable, the basis for our judgments are scrutable, so we have the obligation when we make judgments such as, are these collection -

- is this collection of equipment degradations that happen to have coincide at the same period of time related to a single underlying cause or are they all completely independent of each other, and it was just happenstance that they all happened to line up. You're right. Those require judgment.

Now what we've tried to do is help to provide some guidance for that judgment, so we've

tried to set a threshold above the cross-cutting issue threshold, or above the management threshold, because I think we can all acknowledge that it would be easy to say if you have a collection of degradations that occur concurrent in some period of time, that we can just lump them all into a pot that's called management, you know, deficiency. And we could do that, you know, with whatever issues came up. But what we tried to do is suggest that in keeping with the risk-informed aspect or objective of the program, we try to make a distinction. We try to say if -- and the example that was given in the procedure itself, if you have a bad maintenance procedure and it's applied to a number of different things and they all happen to degrade the same way, it's an easy, that's a fairly easy call to say there was a single issue and it had the effect, the risk impact of the collective multiple degradations, so we enter the action matrix with a single issue.

CO-CHAIRMAN APOSTOLAKIS: I think this is a good example of focusing on performance rather than risk. Say that you have two deficiencies or two problems that may both affect an accident, so they raise this. But you decide that they're due to independent causes so you treat them separately. So

from the risk perspective now you went up, but from the performance perspective you didn't, because they were just random occurrences or whatever, and you're focusing on performance.

MR. COE: Right. First of all, the individual issues would be still inspection findings. There's still a degradation there or deficiency there, and perhaps what you're suggesting is what could occur is that they could both individually treat it independent of each other be green findings. And yet, when you -- because they apply to the same accident sequence, maybe there's a synergism there that causes the collection of those two things to be greater than green. Maybe it's white, or even yellow.

One of the things we captured in the very last sentence in our guidance, if I can essentially paraphrase, that what we're trying to say here is, is that in any case, the staff should be honest and forthright about the collective risk impact. But for the purposes of entry into the action matrix, we may end up with two green issues, which may not prompt additional action. However, if the collective significance of these independent issues was greater, you know, was of some threshold that should prompt some response, that response -- we have tools in our

inspection program to do that, tools that do not depend on whether or not the -- do not depend on decisions like this one as to whether or not we have independent findings or a single finding. And that is guided by Management Directive 8.3 which allows us to initiate a special inspection in AIT, or even an IIT as a response that is risk-informed, so if in fact we have a significant issue here, risk-significant that is dependent upon these multiple equipment degradations, even before we know if there is a common underlying cause or not, we have the tools in our program to initiate an additional inspection to try to get more information so that we can come to a conclusion.

CO-CHAIRMAN APOSTOLAKIS: Now on this issue, it may not be directly related, but I think it is related. When I read the inspection manual on page B-6, which is Appendix B and 7, there is reference to follow-up. "The NRC normally follows up plant events in three ways, events of low safety significance, events of moderate safety, and events of greater safety significance." Later on it says, "Plant status purposes and identifying and understanding emergent plant issues, current equipment problems and ongoing activities and their overall impact on plant risk".

And then later on on the next page it says, "The supplemental element of the inspection program was designed to apply NRC inspection resources either by inspection findings evaluated using the SDP or when PI thresholds are exceeded." Right, B-6 and B-7.

Now there is a lot of focus here on events. Perhaps what we've learned from Davis-Besse is that we should not focus on events alone, that there is -- what if there is information that, you know, like erosion of the head, the vessel and so on and then it was discussed here and so on. It's not an event at a particular plant, but there is this information that is out there. Shouldn't the inspectors take that into account when they decide on supplemental inspections and so on? I mean, maybe that would be a way of handling something like Davis-Besse, not just focus on what's happening at the plant, but also take into account outside information that is relevant to the plant and ask yourself well, are they doing anything about it? Should we have a supplemental inspection regarding this?

You know, that's touching now on the issue of safety culture and so on, but in a more pragmatic and realistic way. What's the difference between realistic and pragmatic? They sound nice. If you say

both of them, they sound nice.

MR. COE: I would agree.

MEMBER SHACK: Realistic could still be unpragmatic. It frequently often is. To be pragmatic you have to be conservative.

CO-CHAIRMAN SIEBER: Which would be this side of the table. Right? Let me expand on that a little bit. If some incident occurs in the Far East. It would be pretty much of a stretch to expect every licensee to be fully informed of that kind of an incident. On the other hand, if there is information that is issued by the NRC on that, or perhaps an INPO SOER or SER, something like that, I think the expectation from the staff would be that the licensee should deal with that. And there used to be an inspection module that followed up to see if licensees actually did review all this information back in the days when you had 100 information notices a year. I would think that it would have to come to the licensee in some official or semi-official kind of way before you could include that as something you would expect them to do or know about in the process of operating the plant. Is that correct or incorrect?

MR. COE: That's correct. And we have tools to do that. Clearly, an issue, an IEP, a

bulletin, information notice as you say, regulatory information summary. But typically for the significant issues, we issue a bulletin which requires a response. We also implement in many of those cases, we implement a temporary inspection instruction to ask our inspectors to go out and review the actions that licensees took in response to that bulletin. And then that instruction is closed out when all of those temporary actions, or all of those inspections have taken place. And then we evaluate whether or not, based on the results of those inspections whether or not we should make other changes to the program.

CO-CHAIRMAN SIEBER: On the other hand, one the staff takes those actions, the licensee is expected to respond to it, and you have the tools to follow it up, and so the fact that some incident in a foreign plant or domestic plant for that matter that bears on a condition in the licensee's plant is relevant and should be part of the ROP.

MR. COE: In some cases we would agree, because we do benefit from operating experience review. In fact, we're taking a very dedicated look, re-look at how this agency in total handles operating experience. And this is motivated in large part because of the lessons learned from Davis-Besse, so we

are taking a look at that. But we have -- my point, I guess, is that we have tools in place now, and I think that we're looking at how to employ those tools more effectively or even to create new tools, if that might be appropriate.

But I'd like to get back to Dr. Apostolakis' point, because it is true that there may be information and circumstances that help provide insight and input to a decision on supplemental inspection activities. But I would offer that the way that that's done is that the initiation of supplemental inspection continues to rest on the objective facts, you know, the performance, the particular degradation that the deficient performance caused. And once that threshold is reached that, you know, we come to the pre-determined conclusion in the action matrix that we would initiate in supplemental inspection, the focus and the specific activities of that inspection. And this is a point that is often lost in these discussions, is informed by everything that the inspectors know to be true or know to be problem areas in the plant that they're inspecting. Our baseline inspection programs day-to-day are informed collectively by the collective understanding that a resident inspector develops on a day-to-day

basis at the site. The samples are chosen within the baseline inspectable areas, you know, with an eye towards trying to identify the most significant problems that may exist at that site.

This is true of a supplemental inspection, as well. The circumstances that prompted the supplemental inspection are certainly clear, as is the history and other, I guess you would put it circumstantial evidence and information that would imply maybe a deeper lying issue, and this forms, you know, part of the basis for how that supplemental inspection is conducted. So I would offer that that's -- you know, I don't know if it satisfies the question but that is how the program deals with that.

CO-CHAIRMAN APOSTOLAKIS: Well, even for Davis-Besse though, I mean maybe it was not part of ROP, but the NRC did ask for extra inspections. I don't know if you want to call them supplemental. It did. And then the argument was, you know, when to do it. Should they do it in March, in February, in January, on December 31st. Right? But that outside the ROP wasn't it?

MEMBER SHACK: Yes.

CO-CHAIRMAN SIEBER: That was a bulletin.

MEMBER SHACK: That was a separate thing.

Right.

MR. COE: Are we talking about after the head degradation issues were revealed?

CO-CHAIRMAN APOSTOLAKIS: No, before.

MR. COE: Before.

CO-CHAIRMAN APOSTOLAKIS: You know, the issue of doing it by December 31st.

MEMBER SHACK: The inspection for the cracking.

CO-CHAIRMAN APOSTOLAKIS: For the cracking.

MR. COE: For the CRDM cracking. The licensee's own inspection.

CO-CHAIRMAN APOSTOLAKIS: Yes.

MR. COE: I understand.

CO-CHAIRMAN APOSTOLAKIS: That was outside the ROP. All I'm saying is perhaps we should think about the language here that this is focused on events or indications that people see at the plant, and those may trigger a supplemental inspection. I'm saying that it may not always be an event, may be some piece of information, and I agree with Mr. Sieber that it has to be transmitted through formal channels. We don't expect those guys to read all the journals that are published, and know everything that happens EBF or

EPCO, but given the risk has happened, unless it's something big like Davis-Besse where the inspection was required now by a different group, shouldn't there be a way here of triggering a supplemental inspection? Maybe that would help us catch the big events. They don't give you advance warning, but then, you know, you realize after the fact that you came close to something really bad.

MR. COE: It's a legitimate question. It's one that's been on our minds from almost the very start. The way that we articulate that question is that there may be issues that a risk evaluation is so difficult to do, to accomplish that the effort it takes to do that, there's -- you start running into a cost benefit issue here. How much -- do you continue to invest money and dollars, and resources and time, and it becomes more and more untimely as times goes on to try to get to some answer. You know, there's a question that's on the table, is there a cost benefit crossover point where we just say it's not beneficial to continue down this path. And there may be another -- there may be a need for creating another mechanism to prompt the supplemental inspection that you're discussing. But the original objectives of the ROP are still very much on our minds. We want it to be an

objective determination, and the more you allow the more subjective elements to enter into that decision process, of course, the further away you get from being objective. And we want it to continue to conform, you know, with understandable, scrutable and repeatable, or you know, consistency in how we treat one licensee to the next. And so all of those -- it's a classic engineering optimization problem. Right? You've got all these competing goals and objectives, and you're trying to find the right balance.

We haven't answered the question that I've articulated, which I hope is similar to the concern that you've expressed. And it is an action item on our SDP Improvement Initiative to resolve that question at some point in time, so that's the best I can give you as an answer right now. It's a good question.

MEMBER SHACK: I mean, since you're so bound and determined to make this process risk-informed, why aren't you adding up the risk from everything that you find and using that as your trigger?

MR. COE: In terms of the concurrent findings that may exist?

MEMBER SHACK: Yeah.

MR. COE: I guess it has to do with having to decide what goes into the action matrix, and all the various combinations of the ways that issues can arise, and what periods of time, and what various individual significances they might have. And we felt that the action matrix, and this is a good point to make. The action matrix in a very high level way aggregates and sums, if you will, the inputs that come to it.

MEMBER SHACK: After I've screened it, and screened, and screened.

MR. COE: After each -- after we've decided that there is individual deficiencies that meet from each other, and therefore, independent. And therefore, their significance characterization is analyzed independently from the others. But then we input those collectively and we have an aggregation. Now that doesn't necessarily catch the synergies that may occur for some specific independent issues that may -- there may be a synergy there from a risk standpoint --

MEMBER SHACK: Synergy, smynergy. Risk is a scaler. It's additive. I just add it all.

MR. COE: And that's --

MEMBER SHACK: Even if I'm only capturing

a portion of it so heck, if it's getting big already, I know I've missed something else.

MR. COE: And that's essentially the philosophy behind the action matrix. But our difficulty from a program office standpoint was in helping to decide when is a collection of issues to be treated independently, and all provide independent inputs to be aggregated in the action matrix, or when should they all be lumped together into a single issue and input into the action matrix as a single finding? So we've tried to give some structure to the decision process acknowledging that it's not cut and dry. I'm not sure I've answered your question.

MEMBER SHACK: Well, you're determined to be risk-informed, that you're only looking at your risk one aspect at a time, you know, until you get to this final action matrix, then you do some combination. You know, if you were risk-informed, I would think you would be looking at the aggregate.

MR. COE: And again, the golden rule here is that we will be honest and forthright about any collective significance that may come from multiple degradations that occur at the same period of time. If we choose to split those up and to be independent, we have to be very clear about that we did that, and

why we're doing it.

CO-CHAIRMAN APOSTOLAKIS: In accident sequences, or even the incidents we've had in the past, you know, there's always a sequence of events. It's not one thing. It's not clear to me that there was a common underlying cause. The valve staying stuck at Three Mile Island, what did that have to do with auxiliary feedwater system being unavailable for eight minutes? It was a different thing, yet a combination of these things led to something. So why then not analyze them as an aggregation of things, rather than looking at the underlying cause? Now again, an argument in the name of performance, you might say yeah, I can look at these things separately, but in the name of risk you have to look at them -- you know, I look at the plant at one instant. This is what I find; therefore, risk is this. I remember somewhat earlier you said that even when they're treated separately, the inspectors are required to actually do the aggregation, as well. Is that still the case?

MR. COE: Well, the inspectors are required to do --

CO-CHAIRMAN APOSTOLAKIS: Or the analyst, the reactor analyst.

MR. COE: Yes.

CO-CHAIRMAN APOSTOLAKIS: Even if you have two events and they are judged to be not to have a common underlying cause.

MR. COE: Correct.

CO-CHAIRMAN APOSTOLAKIS: At some point, there is a risk evaluation considering them concurrent.

MR. COE: Right.

CO-CHAIRMAN APOSTOLAKIS: I thought that was the case.

MR. COE: That is what we tried to capture in the last sentence of the guidance. In all cases the risk of concurrent multiple equipment or functional degradations, and our basis for treating these as either being common cause or being independent should be documented in an inspection report, so we want to be honest and forthright. If there are these -- there's this collection of issues created at a particularly significant period of time, we want that to be very clear. But it goes back to, I think what you were saying, the action matrix deals with the performance of the licensee, and so it wouldn't necessarily be appropriate if there were two completely, at least in our way of thinking, have two

completely independent performance deficiencies that could have happened at any period of time, but just probabilistically happened to happen at the same time. And in many cases, I'm not sure that there would be a real significant difference in our action or our response. At least the evidence to date suggests that there wouldn't necessarily be a difference in our response, that the combination of those two things, whether we call them a single issue or two independent issues, that we would have much of a different response.

CO-CHAIRMAN APOSTOLAKIS: It's still not clear to me, coming back to Davis-Besse, there were indications like the air filters, containment and so on, did the inspector supply this thinking there?

MR. COE: Well, I don't think so, only because the issues at Davis-Besse arose over a period of time which span both the old program and the new. And the type of thinking that you're suggesting is appropriate, we would suggest is appropriate also, to think of or to observe how various indications of degradation could potentially combine together to be particularly risk-significant. And we would hope that over time as the program provides the tools for inspectors to become more risk-informed, that they

would be more sensitive to things like that.

I'm not sure how well that applies to the specifics of Davis-Besse, because I mean, the fact that they saw some coolers clogging up in the containment, you still have to make the logical connection to where that material is coming from, and that it could potentially have come from the reactor coolant pressure boundary.

CO-CHAIRMAN APOSTOLAKIS: That's good thinking. Why didn't they ask that question? Why didn't they say that? Because the moment you say that, I mean, maybe risk analysis would tell you that boy, we better look into it.

MR. COE: Exactly. And that comes from a sensitivity to what could potentially be the significance of a degraded reactor coolant pressure boundary. Again, I would hope that over time our inspectors, given the tools and the training that we believe are appropriate, will come to a greater sensitivity of issues that could be -- I mean, the whole program -- the whole reason we have significance determination processes that are publicly available, laid out in a document for our inspectors, as well as the licensee, as well as the public to see is to provide a road map, a yardstick, if you will, of what

things are more significant than others. And we would hope that our inspectors take those road maps and use them to lead them in the areas of greater significance, and to help them differentiate the things that they probably don't need to pay as much attention to from the things they really should.

CO-CHAIRMAN APOSTOLAKIS: Is there an investigation going on now, how the -- what lessons the ROP should learn from Davis-Besse? Is that what -

-

MR. COE: Absolutely, yes. Yes, sir. We have -- in fact, we're well along in that process, and have been given the results of a very substantial task force effort that has specific line items that have been handed to the program office for direct oversight process, and that we're taking a very specific look at. And it involves utilizing operating experience better, improving our ISI inspection procedure, and a host of other things in terms of operator sens -- I'm sorry, inspector sensitivity and training.

CO-CHAIRMAN APOSTOLAKIS: When do you think this Committee will find out about this?

MR. COE: I don't know the answer to that.

I'm asking for some help.

MS. CARPENTER: Okay. I think the Davis-Besse task group report is issued. And then what the staff is now doing is taking all those recommendations and we're putting them into action plans. And we have a due date to the Commission with those action plans, I think February 28th.

CO-CHAIRMAN APOSTOLAKIS: This February?

MS. CARPENTER: This February, right. That's the action plan on all the items that we need to do, and we're starting to work on those now.

MS. WESTON: That's just the action plan.

MS. CARPENTER: That's just the action plan.

MS. WESTON: Not the responses to the issues raised in the action plan.

CO-CHAIRMAN SIEBER: And you're referring to the Lessons Learned Task Force.

MS. CARPENTER: Right. The Lessons Learned Task Force.

MS. WESTON: Right.

CO-CHAIRMAN SIEBER: There is no other task force.

MS. CARPENTER: No.

CO-CHAIRMAN SIEBER: Other than that.

MS. WESTON: Right.

CO-CHAIRMAN APOSTOLAKIS: But I didn't see there any statement as to why the inspectors at Davis-Besse acted the way they acted. It just says that the NRC failed in certain respects, so how can you learn from that? Anyway, are we going to see this plan?

MS. WESTON: Yes, we will.

MS. CARPENTER: That plan -- my guess is that plan should become public. And the staff is beginning to --

CO-CHAIRMAN APOSTOLAKIS: No, no, no. Not as members of the public. Come on. Are we going to review it?

MS. CARPENTER: I don't know the answer to that.

CO-CHAIRMAN APOSTOLAKIS: I know I'm a member of the public.

MS. WESTON: No, we will put it on the agenda.

CO-CHAIRMAN APOSTOLAKIS: February 28th is too close.

MS. WESTON: No, February 28th you will only --

MS. CARPENTER: Just the action plan.

MS. WESTON: -- is only the action plan. The EDO sent a letter back to research and NRR asking them to do action plans for the issues that came out of the Lessons Learned Task Force.

CO-CHAIRMAN APOSTOLAKIS: Okay.

MS. WESTON: So the only thing that they're going to do there is to say this is our plan to address those issues. There will be no issues addressed in the February 28th --

CO-CHAIRMAN APOSTOLAKIS: And we will be briefed after the issues are addressed?

MS. WESTON: When -- as the issues are being addressed, hopefully.

MEMBER KRESS: I guess George's concern is, have they identified the right issues.

CO-CHAIRMAN APOSTOLAKIS: The right issues, and also, you know --

MS. WESTON: Well, you have to look in the Lessons Learned Task Force for that. The 50 some recommendations in there are those that -- with the exception of two I think went forward. The Management Task Force recommended that they look at all of those issues with the exception of two.

CO-CHAIRMAN APOSTOLAKIS: It's not just the issues though. It's also what you plan to do

about it. I mean, everybody keeps raising the issue of questioning attitude, but what to do about it is a monumental problem.

MS. WESTON: That's what the action plan is supposed to address.

CO-CHAIRMAN APOSTOLAKIS: Okay. And that's where I think we should --

MS. CARPENTER: In reality, the action plans are only addressing the high priority items. There are a number of items that were medium and low priority, that many of the branches are already beginning work on, that we're taking them -- we have to make sure we budget the resources and everything into these. So those action plans will only address the high priority items. I'm thinking there are about 28 of those. But there are a lot of others that will simply be put into our budget, and we're going to start working on them. We are starting to work on.

MR. COE: And they go well beyond our program.

MS. CARPENTER: Exactly.

MR. COE: I mean, ROP is a part of it, but not the whole picture.

CO-CHAIRMAN SIEBER: Well, I think fundamentally you're getting down to who has what role

in the process of operating the plant. The resident inspectors, there's usually a couple of them at the plant, and a stenographer, and two people cannot cover every minute aspect of the operation of the plant. And so whether a filter clogs up some place, which is not a regulatory event typically. That's just something that is a maintenance or a service item, whether that clogs up in conjunction to the fact that somebody issued the bulletin, and you may have suspected CRDM cracking, I'm not sure that one would expect the ROP or even the resident inspector in his normal function to be able to put all this stuff together to say to the licensee, I think you have a leak and your reactor vessel head is degrading.

I can see if he were qualified as an operator on that plant, and that was his responsibility, like operators are supposed to have it as their responsibility, then he could put it together because that's what operators do, and it's the licensee who is supposed to operate the plant. And the NRC is supposed to regulate how that plant is operated. In other words, are all the programs and processes in place. To me, I think there's a little bit of confusion as to whether there should be an expectation on our part that the resident inspector

should have deduced the fact that the head was being degraded.

MEMBER BONACA: I agree with you, and I think that, you know, hopefully this task force also looks at the issue of whole strategy accepted by the NRC for the CRDM cracking was the visual inspections. And yet, at Davis-Besse the three top nozzles were never inspected visually. They never accessed them, so here you have a situation where we are setting up for failure really plant personnel at the working level, because they don't set the strategy as well as the resident inspector and everybody else, by the fact that a fundamental requirement to support the strategy of just depending on visual inspections has not been implemented, and has not been followed through. And I'm not sure that I read that in the root cause, but I think I can read it through some of the recommendations, but it's not so explicitly stated.

MS. WESTON: With regards to the resident inspector, that issue was discussed at length at the Commission meeting last Wednesday on the Lessons Learned Task Force, and the commissioners raised several questions regarding what they would do about the resident inspector and their learning process in terms of being able to raise issues that they were not

capable of handling themselves outside of that process. And with regards to the inspection part of it, NRR is, in fact, taking another look at the inspections at the plants. And we will hear more about that later.

MEMBER BONACA: I know that now. I mean, everybody is asking why we would do what the French did. But, I mean, you know, that's however -- the bigger issue is even though we had a different strategy, why didn't we follow through by assuring that in fact the inspection would take place? And after 12 years or 10 years from the first finding of this cracking, still those three top nozzles were never looked at. I mean, that's a pretty significant issue that sets up everybody else for failure, you now, including, of course, the resident inspector who is the guy who is not going to go up there and look at it himself. He again is doing other things, and he failed. Maybe we failed.

CO-CHAIRMAN APOSTOLAKIS: Another issue we have raised in the past is this assumption that if the safety conscious work environment has deteriorated, we would see that in equipment performance. And I see th is inspection manual repeating that. It says, "In short, no separate and distinct assessment of

licensee's safety culture is needed because it is subsumed by either the PIs or baseline inspection activities." And it's not even dated, so I presume it's still draft. Should we really say things like that now in light of what happened at Davis-Besse? Shouldn't we just soften it a little bit and say that maybe we are thinking about it, and what to do? Because clearly, that's not the case.

MEMBER ROSEN: Where are you reading, George? What page?

CO-CHAIRMAN APOSTOLAKIS: Page 11 of the -

-

MR. COE: I think you're reading the draft basis document.

CO-CHAIRMAN SIEBER: Right.

CO-CHAIRMAN APOSTOLAKIS: Well, it says, "NRC Inspection Manual, Chapter XXXX."

MR. COE: That's the draft basis document.

MS. WESTON: It's the basis document.

CO-CHAIRMAN APOSTOLAKIS: So what does that mean?

MR. COE: It means it provides the basis for the rational and the basis for how we --

CO-CHAIRMAN APOSTOLAKIS: The very last

sentence of this section, "Safety Conscious Work Environment", repeats this assumption.

MR. COE: Yes.

CO-CHAIRMAN APOSTOLAKIS: And, you know, we questioned it in the past, and I wonder whether after Davis-Besse we should still say that.

MR. COE: That is a subject that's on the table for us to examine.

CO-CHAIRMAN APOSTOLAKIS: So why don't you say that, that we are thinking about that. I mean, that was a prior assumption, now we are --

MR. COE: Basically because the basis document represents the current philosophy, the current basis for the current program. We are saying that we're -- you know, pursuant to our effort to, you know, respond to the task force on Davis-Besse, we are going to look at this. But I can tell you that early in the program, you know, it made sense that if you had cross-cutting issue problems at a plant, that they would over time reveal themselves, and we expected to pick up on those manifestations of that underlying problem.

There was some thought given to how to inspect safety conscious work environment directly through the use of survey instruments, such as the one

that the Office of IG utilized for the NRC staff, but that was dropped from further consideration principally because of the cost involved of exercising that kind of an instrument, you know, at our licensees over a period of time.

CO-CHAIRMAN APOSTOLAKIS: Its value is questioned. I mean, if you go and ask somebody, do you have a questioning attitude, what is he going to say? No, I'm stupid, I never ask questions. Come on. This is ridiculous. These surveys don't mean much in my book, but coming back to your point though. I think in many cases you're right, there will be deterioration that will be observable some place.

Unfortunately, there are some cases, for example, involving barrier integrity like Davis-Besse, where you may not have this luxury of advance warning, and this where, you know, we may want to do something about it, but I don't think it's something that can be resolved in a week or in a month. But I was just struck by the statement. I mean, it's as if nothing has happened. I mean, I know that this was the position in the past, but I would expect it to be softened by now. In fact, there is another statement up there, possible indications of an unhealthy safety culture include a high number of allegations of weak

employee concerns program and a high corrective maintenance backlog. None of these would have caught Davis-Besse, so drop it then. You don't need that. Why do you have people like me criticizing that? I mean, Davis-Besse is not -- I think it was a major test of the ROP, and I don't think that you gentlemen and lady think that way. I thought it was a major test and it failed, and we have to do something about it. And that's why it bothers me when I see these things. I always -- my mind goes there and I say well gee, a weak employee concerns program. That has nothing to do with Davis-Besse.

MR. FRYE: This is Tim Frye from the staff. Doug already mentioned this, but I just wanted to re-emphasize that the basis document is trying to capture the basis of the program as it exists today. And it's also important to remember that it's a living document. And that as the ROP changes, we'll be looking to update the basis document to reflect the changes we made. But, you know, right now that's the basis of the current program. That's why it reads that way.

MR. COE: I would offer -- I'm not sure that the staff has yet concluded or will conclude that the ROP was a failure with respect to its application

at Davis-Besse. You know, even the Commission acknowledges that the issues at Davis-Besse occurred, you know, well before the ROP came into effect, and the previous program didn't necessarily identify this underlying issue.

The ROP does have tools. Now again, we are taking a look at how to improve the tools based on our lessons learned from Davis-Besse, but currently cross-cutting issues or cross-cutting aspects of inspection findings are captured in a specific place in our inspection reports. And those are accumulated and then made available for the express purpose of making available to our team inspections that look at problem identification and resolution programs. So in addition, we have an opportunity to discuss cross-cutting issues with licensees at our mid-cycle and end-of-cycle letters that we -- assessment letters that we provide to the licensee every six months, so these are the tools that exist, and perhaps we can use them better. Perhaps there can be other tools that we can conceive of that would help in this area.

MEMBER BONACA: I think, however, it seems to me that we are looking at, you know, safety culture and the stop gaps to a situation that had other elements in it. And I brought up already one before,

this lack of follow through. The whole industry, I mean, we're really -- the way it was handled, the CRDM cracking, the leakage, the assumptions that, you know, boric acid would not corrode, evidently, you know, the carbon steel. I mean, because of all the reasons, set up the whole situation that cascaded in this. Now then we're looking at safety culture as a stop gap situation that will identify all these problems. Clearly, it didn't. The only thing that surprised me about davis-Besse is that there were no differing opinion, that nobody raised this issue of concern about clogging of the filters. It's almost like, you know, for me there is organization walking lockstep, and everybody had this full agreement on where it was coming from. But, you know, I think that there was a lot of situations that could have been recognized well before that. If you look at the failure of the program, and I'm not that what is taking place is going to really identify that. And that to me, that's really the root cause of the whole thing. Okay. Again, the cracking of this -- I mean, if you think about the whole process that was brought to bear and everybody accepted, and dangerously, and set up all those Davis-Besse people.

MR. COE: And I think that both the

industry and the NRC are owning up to their share of the responsibility here, and that's reflected in the task force's recommendations.

MEMBER BONACA: I'm not looking about the past. I'm looking at what we need to do in the future to prevent situations like this from occurring. You know, because I mean there was a clear distinction there between the way that the French went about it, which was automatic inspection from day one, and that resulted cascading into replacement of the heads very quickly because it's too expensive to do automatic inspections, from the way we did it here, we said we are going to accept visual inspections. And then we didn't eve put forth requirements to have proof that these inspections were being done.

MR. COE: You'll have a opportunity, as was mentioned, to review the task or the action plan to respond to the task force's recommendations. And I think it would be very useful to get your insights on that task action plan to ensure that we are covering all of the aspects that are important. You're making some good points. I don't dispute that.

MEMBER LEITCH: Doug, could I go back to something that you said a few minutes ago, at least I understood you to say. I'm not sure I heard you

correctly; that in common underlying causes, you kind of limit how far you drill down, I guess because if you drill down far enough you can find a common cause for almost anything. But by limiting how far you drill down, don't you eliminate the potential to find some of these cross-cutting issues, like safety culture or management issues. It seems to me that you have to be sensitive in that limitation because if you don't find those kind of -- if you limit your look so that you don't look deep enough to find those kind of cross-cutting issues, it seems to me that you prevent the ability to find some of these safety culture management kind of issues.

MR. COE: It's a good question, and the response is that although that we tried to set a limit on how far you can drill down, as you say, that's a decision result. Getting to that decision, I think intrinsically means that you have to examine deeper issues to try to come to the decision point. Is it or is it not a cross-cutting issue? Is it a common issue that we just lump everything together and call it one issue? So you have to drill down deep enough in order to make those judgments, and I would offer also that our inspectors and our -- and their supervisors and their managers continue to be sensitive to extent of

condition questions, and cross-cutting issue questions, and again, we have elements in our program that they can avail themselves of to document those kinds of issues, and roll them up over time, and talk to the licensee in assessment letters, and use those insights in our PI&R Inspection. So it's a very good question, and those are the ways that the program intends to try to deal with that.

MS. WESTON: Now, Doug, as a cross-cutting issue, is any thought being given to documenting or capturing the number of times that action items are entered into from the tech specs? That's one of the issues with Davis-Besse also. They apparently entered tech spec many times. Is that going to be considered as a cross-cutting issue when you look at the impact that Davis-Besse may have had on the ROP, or what you need to do about changes to the ROP?

MR. COE: Right. If there are issues that keep recurring, obviously in our corrective action program, for example, the equipment, you continue to have to enter tech spec action statement, you know, repeatedly over a period of time because of some deficiency, or failure, or need to remedy some problem. Then those kinds of issues are good sampling opportunities for the PI&R Inspection. And all I can

say is that the PI&R Inspection affords us that opportunity, and I believe it has specific guidance in it that suggests that we look for those kinds of things. A lot of those kinds of things come from our insights that the residents gain over a period of time. They know that certain things are problems. They know that they reside in certain areas of the licensee's plant or their organization. Those are all inputs that are utilized and are useful to picking the samples that we pick. I mean, there's only so much time that you have in these inspections, and so you have to make the most effective use of that time, so we try to pick smart samples and use all available information.

CO-CHAIRMAN SIEBER: On the other hand, you could find a fair number of indicators that would tell you maybe the safety culture here isn't very good. On the other hand, the mitigating equipment operates, meets its test requirements. You don't have a lot of initiating events, and the licensee seems to be getting by. If that's the case, then what does the staff, including the resident inspector do with this new insight they have given their inability to connect the dots, would be the phrase we've heard over the last few weeks.

MR. COE: And that's a good question because we set a higher standard in the ROP for connecting the dots. Clearly, the -- and even prior to ROP, as a senior resident inspector, my inspectors would often come to me with issues that, you know, were not necessarily the smoking guns. It's a feeling. It's like, you know, I think there's a problem here.

CO-CHAIRMAN SIEBER: Right.

MR. COE: And so well, where's the evidence? Okay. I mean, I can't go to the exit meeting with the licensee and lay down that I have a feeling that there's a problem here. Even before ROP we set a higher standard for ourselves. Now in ROP, we not only have to have the deficiency identified, performance deficiency identified, but if it's -- you know, if we're going to take further action in terms of additional supplemental inspection, it has to reach a certain threshold that we've pre-defined. So it goes without saying that I think we set a higher standard for ourselves, but knowing what we know as we walk through the plant day-to-day, day in and day out, gives us clues. And I can tell you, I have a deep affinity that our inspectors face on a day-to-day basis going into these plants confronting enormous

information, quantities of information. I mean, everything they see, everything they hear, every discussion they have with a licensee staff person provides clues, and those clues are the things that they have to pull together and connect the dots with. And it's a very difficult challenge.

CO-CHAIRMAN SIEBER: Well, it's not only connecting the dots. Maybe you firm up your suspicion to some extent, but you don't find a violation of the rules, and you don't find a risk significant situation. I read a speech, as we all did I think, about safety consciousness, safety culture which is different than safety consciousness, is becoming an issue because of Davis-Besse. The question is, should you regulate it, and if you should, how do you regulate it? And I think that that's a very, very difficult problem that's been around since the mid-1970s, and attacked and backed-off of, the subject of negotiations between the industry and the NRC, and all kinds of things. That's where INPO came from, so on the other hand, I see it raising its head again.

MR. COE: And our colleagues in other countries often take different approaches to the direct observation, inspection, and in some cases regulation of those kinds of elements, more subjective

safety conscious management type of issues. And we acknowledge that there's different ways of going. Recall again though that the ROP was driven by a desire to be more objective and move away from that because it was perceived by some as having given us too much latitude, and it was not being consistently applied.

MEMBER ROSEN: I would add to that discussion that safety culture in aviation and medicine has been recognized as a prime determinant, and I happen to be holding in my hand a book by Helmreich and Merritt called Culture at Work in Aviation and Medicine, which talks a lot about how the aviation industry particularly came to the conclusion 20 years ago that the culture of the cockpit, crew resource management is important whether or not people got to their destination site, so now we're faced again with the same discussion.

CO-CHAIRMAN APOSTOLAKIS: We will have this some other time.

MR. COE: Lessons to learn there.

CO-CHAIRMAN APOSTOLAKIS: Coming back to the ROP.

MR. COE: Yes.

CO-CHAIRMAN APOSTOLAKIS: This basis

document again.

MR. COE: Yes.

CO-CHAIRMAN APOSTOLAKIS: Page C-3, "Use of change in core damage frequency versus condition of core damage probability." I'm not quite sure I follow this. You have a number of findings, and you decide there is a common underlying cause, let's say. Then you have a choice whether you want to calculate the CDF or CCDP. Is that -- that's what this says. "The SDP can be used to estimate either CDF or the CCDP given any plant configuration, which may include the combination of degraded equipment functions and equipment outages for maintenance." And then you say, "The staff recommends the use of the estimated change in CDF instead of CCDP." And I'm trying to understand what does that mean?

MR. COE: The choice of using the change in core damage frequency is derived from the need to have a baseline core damage frequency that we accept to be -- that we accept as acceptable that includes periodic maintenance, et cetera, and over time there's actual -- day-to-day there's a change, but on average there's a baseline core damage frequency that includes maintenance activities and other testing activities, that sort of thing. And that what we're trying to do

is to measure the impact the licensee's performance deficiencies had on public health and safety risk by choosing a metric that is the increment of risk above that baseline, that nominal baseline.

The way that's done is to take the CCDP for the particular time period involved, and then that CCDP is essentially normalized across the entire -- on a per year basis to be compared to that nominal baseline CDF, and then increment then is a delta CDF.

CO-CHAIRMAN APOSTOLAKIS: Is that the same as saying I found that the unavailability of two systems was higher than should be for a period of say three weeks? Now what is the probability of having the initiating event in that period? Because if I have an initiating event, then I'm in trouble, so that would be the CCDP. But then you normalize it over the year because it was three weeks only. I mean what -- does the ASP do th same thing? The ASP calculates --

MR. COE: It does -- in essence it does use the same metric. Although the -- remember that we're also making a distinction between an event and a condition. An event is always evaluated in terms of the probability of core damage given that the event occurred. And a condition involves all the range of possible initiating events that may have occurred

during the time period the condition existed. But that's not the way that I used CCDP just a moment ago. What I used a moment ago was essentially condition CCDP. I'm going to take events off the table for the moment, because the SDP deals with conditions.

CO-CHAIRMAN APOSTOLAKIS: Are any of your examples later involving conditions versus events, because we have a number of them.

MR. COE: All of the examples are the reactor safety SDP involved conditions, because that's the only thing that the SDP analyzes for use by the action matrix, is conditions. The moment in time, probability of core damage when a particular event may have happened is a metric of interest to us, and the ASP Program will, in fact, attempt to evaluate that. But it's not considered an input to the licensee's performance, unless we can identify a particular performance deficiency which resulted in some degradation to the plant's design or function that has contributed to that event.

CO-CHAIRMAN APOSTOLAKIS: But this section, I must say, I don't understand. If CCDP were used to characterize licensee performance, the result would be inconsistent as it is influenced as much by timing, that is plant configuration, as by deficient

performance. How can you avoid the issue of timing? I mean, if the condition existed for a week, that's different from a condition that exists for six months.

MR. COE: That is taking into account the time, not the timing. What was meant by "timing" is, if the timing of the deficiency happened to occur at precisely the time that acceptable on-line maintenance was occurring, the procedures in the SDP require that you not include the unavailability of the equipment that was -- which acceptable on-line maintenance was being performed, because going back to what I said a moment ago, the baseline nominal CDF includes all of the -- probabilistically includes all of the maintenance activities that go on over the year, so that all the maintenance activities are normalized to the nominal baseline CDF. And what we're trying to measure is that increment that is due just to the performance deficiency, and not due to the fact that it happened to have occurred when on-line maintenance was occurring.

There is a further mathematical treatment of this particular point after this discussion, I think, if we retained that. We did at one time.

CO-CHAIRMAN APOSTOLAKIS: I didn't see it.

I'm sure the inspectors or the reactor analysts don't calculate delta CDF based on this guidance. This is just a description of what's going on.

MR. COE: This is the basis. The guidance basically says that if you have your deficiency during a period of time of on-line maintenance, that you disregard the fact that the on-line maintenance was occurring. You only evaluate -- you evaluate the increment of health risk --

CO-CHAIRMAN APOSTOLAKIS: Because you're focusing on performance.

MR. COE: Yes.

CO-CHAIRMAN APOSTOLAKIS: You're distinguishing it from risk. You are. You can't do both.

MR. COE: We're focusing on the performance aspect, and how that performance aspect has contributed an incremental additional risk above and beyond the nominal acceptable baseline risk for the plant.

CO-CHAIRMAN APOSTOLAKIS: But in order to do the risk part, you have to consider the fact that it happened during some preventive maintenance period. I mean, you know --

MEMBER SHACK: What you're saying -- I

mean, when the guys goes to count the systems that he's got available, he still counts the system that is having preventive maintenance being put on it.

MR. COE: That's correct.

MEMBER SHACK: That's what you're really saying. Although in the real world, it was not available, because he's looking at performance.

CO-CHAIRMAN APOSTOLAKIS: But Doug says no, I'm also looking at this. Well, you are in some sense, but it's distorted.

MR. COE: If we allowed that to enter into the SDP calculation, the on-line maintenance additional impact to the risk for that period of time, then the outcome of the SDP would be as much a function of the particular happenstance of when the degradation occurred due to deficiency, as it did on the deficiency itself. It would be an influence on the probabilistic timing of that event or condition.

CO-CHAIRMAN APOSTOLAKIS: So the risk part should be affected by that. The performance part should not.

MR. COE: And again --

CO-CHAIRMAN APOSTOLAKIS: The performance has nothing to do with it, because this is --

MR. COE: Uh-huh.

CO-CHAIRMAN APOSTOLAKIS: It could have happened some other place. But it seems to me that you are really focusing throughout this on performance, and you are using risk to do certain things, but you're really focusing on performance, which I think is appropriate. It's appropriate. It's just that some of the stuff on risk is not too solid.

Like this Paragraph A here, "The reactor safety cornerstone performance indicator thresholds were based on the increase in annualized CDF. Thus, in comparing and adding the effects of PIs and inspection findings within the action matrix, is it necessary to use the same risk metric."

In other words, we use risk to define the thresholds for the PIs, and now we have the SDP results of the risk. And because both of them are based on CDF we can add them, although you don't really add them. You consider them as a --

MR. COE: Yes, that's correct.

CO-CHAIRMAN APOSTOLAKIS: I don't have anything else.

CO-CHAIRMAN SIEBER: Why don't we move on.

MR. COE: Yeah. We've gotten a little bit away from I think the earlier discussion on concurrent multiple equipment. And I think that unless there's

other questions about that, I think it would be useful to move to the example.

CO-CHAIRMAN APOSTOLAKIS: But, Doug, again the consistency of it all. The PIs you take a frequency of one initiating event and you change is so much so that you will see a change in CDF.

MR. COE: Independent of any other changes.

CO-CHAIRMAN APOSTOLAKIS: Right. Then you go to the SDP. Now you have a set of findings, and some things happened to occur during a preventive maintenance activity. And now you say no, I'm not going to estimate risk based on what I see. I'm going to assume that this equipment that's under preventive maintenance is available, so I'm distorting the risk assessment. Why does that make sense?

MR. COE: I would say that there is a consistency aspect between the safety system unavailability PI which measures the unavailability of mitigation equipment to how the SDP would evaluate that. And the SSUPI did use delta CDF because its thresholds were set using a representative sample of some risk models that included baseline maintenance. In other words, nominal amount of maintenance, so what you're trying to do is you're trying to set a

threshold for unavailability that has some basis in an increment of the risk that's over and above a nominal plant risk. And we accept that nominal plant risk has some maintenance activity going on during the year, and so that concept carries over to the SDP in the discussion we just had. So I would say that there is a measure of consistency there, and that was the intent. We can explore this further later on.

CO-CHAIRMAN APOSTOLAKIS: But that unavailability will be averaged over the year, you know, just to make a simple example. If you have a two-train system and one is down for maintenance, and now you have an activity that disabled the other train. Okay? And you have no system whatsoever left, both redundancies are gone. I mean, it seems to me that to simply assume, you know, the average unavailability of the first system over the year doesn't measure the significance of the event.

MR. COE: And that's -- again, I'll return to kind of the golden rule here. We're going to be honest and forthright about the impact on risk of that that you've described. And we have tools, such as the initiation of a special inspection, and augmented inspection team or an IIT, that deals with the specific risk as best we can determine it or estimate

it in the early stages as an issue comes up. And we have and we will continue to use those tools to engage additional inspection resources to get to the bottom of what's really going on, because your assumption was that you were doing just acceptable maintenance on one train, and then you had a deficiency that causes the other train to become disabled. And one of the questions is, is that assumption correct? Is that maintenance being done? Is that being done just because of preventive reasons, or is there really some other reason that it's being done?

We need the full facts. We need to get the whole picture, and then we make decisions about whether the causes were related or not related, and then we choose how to input them to the action matrix.

MEMBER BONACA: Because, I mean, it seems to me also you have other considerations such as, for example, in the original design these plants were not supposed to be maintained half-power, and so therefore, you really have set up a system of tolerance of that situation, provided that you have, in fact, a risk evaluation done ahead of time. And we talked about that. I mean, if you have multiple systems out of service, and in fact -- and that the licensee takes care of protecting the redundant train,

and doing all those things that need to be done when you're taking a system out of service, so you have a lot of considerations you have to take care of. There is a lot of responsibility of the licensee taking a train out of service for maintenance.

MR. COE: Indeed there is, and we've acknowledged that and provided the maintenance rule to set some standards so that the licensee can perform this kind of maintenance because we acknowledge that there can be a benefit, a safety benefit from the performance of that kind of maintenance, and so we've accounted for that via the maintenance rule, and we account for it in the SDP by allowance of it, and such that it does not affect our evaluation of the risk impact on the public, when we're really after that increment that was due specifically to that performance deficiency, and not due to anything else.

Okay. Mr. Chairman, move on?

CO-CHAIRMAN SIEBER: Yes.

MR. COE: Okay. Actually, I've got about an hour to cover several examples, and we'll go through these at a high level, but we'll try to get to whatever level of detail you're interested in. And I do have the detailed packages that the SERP panels looked and reviewed, in case that I don't have enough

detailing in the slides.

I think what -- again going back to the success objectives of the meetings, is to go through these examples, not just -- we're going to start with the reactor safety cornerstones, and then subsequently hit the other cornerstones. But the idea here is to give you a feeling for where our thresholds are in terms of our response to, you know, how we respond to a red finding, or yellow findings across cornerstones, or white findings, so we've given some examples in here of red findings and yellow findings that I'm going to speak to, but I think it might be perhaps more informative for you to consider that the thresholds for all of the white findings, you know, collectively to give you a sense of where you think -- whether you think the thresholds are about right or not for the level of response that we're giving it. And, of course, we're always interested in your thoughts and insights on that.

To start with, the first example that we have here is Example A. It starts on page 11. The issue here was essential service water pump that failed a surveillance flow test, and it was determined that the licensee had allowed some Tygon tubing to enter the intake bay and become lodged in the

impeller. The essential service water system for this particular plant provided cooling water to the diesel generators, contained coolers, CCW heat exchangers, a number of other -- it's a safety-related system so this condition based on their evaluation existed for approximately 132 hours, or about seven and a half days.

The issue screened through the SDP logic and resulted in a white. And if you'll turn to the next page, that presented at a high level. First of all, the Phase 1 screening logic was to ask the question, does it represent an actual loss of safety function for a single train greater than an allowed outage time? And if that's true, then a further analysis is required. In other words, we can't -- there's a potential for it to be greater than green and, therefore, we want to do some further review.

When the Phase 2 analysis was applied, what was identified was that of all of the sequences that this deficiency or this degradation affected, the one that was most dominating in terms of a risk evaluation was the loss of off-site power sequence, that essentially represents a station blackout that persists for up to five hours. At that point, the assumption is made that the core will become damaged

as a result of the loss of seal cooling, and the resultant loss of coolant without recovering any power. So in the Phase 2 level of detail, the loss of off-site power frequency was given a value of three, which represents ten to the minus three, and that was a combination of the time; that is the 182 hours, and the expected return rate or frequency of loss of off-site power. So three represents ten to the minus three essentially and higher, an order of magnitude higher.

The emergency AC power is represented as ten to the minus two, and that reflects the fact that one of the trains of emergency AC power, that is one of the diesels is rendered inoperable because of this particular deficiency, such that if the loss of off-site power occurred, this particular emergency service, or essential service water pump that feeds one of the diesel generators would also -- is already disabled. Therefore, there's no -- it would not support that particular diesel generator. So, therefore, there's only one diesel generator left in a loss of off-site power scenario.

And then finally, the recovery of AC power does not occur within the five hour time period, it is given a likelihood or probability of ten to the minus

one. That's represented by the one. If you add these figures up, and these again represent the negative logarithm of the actual values of probability being used, you get three for the loss of off-site power probability during that period of time, two for the failure of the one remaining diesel generator, and one for the loss or the failure to recover within five hours. That represents a total of six, or ten to the minus six, which represents the low end of a band that represents the white significance level.

Now I will say that as in many cases, this is a way of just talking about the influences, various influences and assumptions that are built into the staff's determination of the significance, the color of the significance. In this case, as well as many others, we do additional analysis with detailed computer-based models. The licensee does analyses and so forth. I this particular example, those analyses supported this result, and I'll just leave it at that.

The dominating influences were similar, and so we could rely on a computer-based model. We could rely on the licensee's model, but none of those -- neither of those would be as scrutable as the representation that's given here in a very summarized form.

MEMBER LEITCH: Suppose the Tygon tubing had the potential to affect the other emergency service water pump, how would that change the analysis? Would that make the actual loss of safety function a higher number?

MR. COE: Without knowing all of the details, what I've read in the package would suggest that there was a single Tygon tube attached to a funnel that was being used near the intake structure for this particular train. This happens to be the B Train that was affected. Without knowing the plant's arrangement and design, I'm going to somewhat speculate that there was only the one tube, and it was only going to go to that one pump. And if that's the case, then there would not be a common cause influence. But if there is a common intake structure --

MEMBER LEITCH: That's my question, there had been a common intake structure.

MR. COE: Right. If there was a common intake structure, you know, with one Tygon tube, it would be expected to have impacted only one of the two pumps perhaps if there's a two train system. I know that, and if there's only one essential service water pump in each train, then it would only impact one in

any case. If there was a difference; that is, if it could have impacted the other one, and that would have potentially represented greater significance because of the loads that it served, then there might be a difference. I don't know that that was the case.

MEMBER LEITCH: Yeah, I understand. I was just wondering --

MEMBER SHACK: You would have lost the two, and so you would have been a three plus one, four. You'd have been very bad news.

MR. COE: It could have been worse if it had been another pump that had a greater -- you know, had greater loads on it, or could have been potentially more significant to have lost that. But in fact, what we are going to evaluate though still is the actual degradation that actually occurred, and that's a given. And the fact that that occurred represents that loss of function for that period of time. And then we look at all of the various initiating events that could have happened during that period of time. And in this case, it was the loss of off-site power that came up as the one of greatest significance.

In a more detailed evaluation, you would have summed up all of the other sequences of lesser

significance, and you would have done that in a very complete way, and so that's what the computer does very well.

MEMBER WALLIS: It may be that there are many sequences which are equally important. When you add them up, you get a different answer than if you just look at the --

MR. COE: That's correct. But in this particular case, you know, checked against the other more detailed models, risk models, both the licensee and the NRC concluded that this was, in fact, a dominant influence. But it does only represent that Pump B was the one that was affected, and whether or not it should have represented that there was a possibility the other pump was -- could have been affected, I don't know, but it wouldn't have made any difference in any case, if only one pump could be affected, and pumps were equal in all respects.

MEMBER KRESS: I have a couple of questions about this. One of them is, does this necessarily represent a poor performance of that particular licensee? That's question number one. Question number two is, suppose this licensee was South Texas, and they had a CDF of ten to the minus whatever, and its role -- a performance role of the

licensee could be viewed as to keep their CDF below an acceptable value. Now with South Texas, this thing would not even have gotten anywhere close to an accepted value, for some other plant though that might have exceeded an acceptable value. So the two questions I have is, should we treat this differently as a plant-specific issue? It should be different at different plants rather than look at the delta. If you had looked at the actual absolute value of CDF, which would incorporate all those other things. And in my view, what should the plant have done differently that would have been better performance? I mean, is this really necessarily a bad -- an indication of a bad performing plant?

MR. COE: Well, to answer your second question first, Dr. Kress, the performance deficiency was noted to be that there was no procedure for installing or removing the temporary drain hoses, and that there was a lack of a questioning attitude the length and the duration of the event. They had several opportunities to question the location of the Tygon hose and failed to do so.

This is a judgment, the fact that this is being viewed as a performance deficiency, but the staff's basis has been identified in the inspection

report.

MEMBER KRESS: I would have automatically given those things some sort of performance criteria that would have given probably more than white, rather than going back to a CDF and --

MR. COE: That gets to your second question, or the first question which I'll now answer. And that is, would it be appropriate to represent the significance on the basis of some absolute risk value? And the choice that was made in this program is to evaluate the licensee against their own nominal baseline risk level that we believe is acceptable. And it's acceptable if you assume that all of the plant's design features are available, given that there's some likelihood they might not perform when called upon, and that's reflected in the probabilistic values of failure probabilities and unavailabilities that we apply in a risk model. So given that, each plant is judged against it's own - and I think that was a question that came up earlier - as against its own licensing basis essentially. And that was felt to be more fair, I guess, if you will than to try to hold every plant to the same absolute standard when all plants are designed with differences. And there might be a range of acceptable risk, nominal risk values

that are still acceptable depending on which plant you go to.

MEMBER KRESS: I think this is another reason that I would like to see the system divorced from risk considerations, and actually be performance-based.

MEMBER ROSEN: If you had a third train here so that this thing wasn't risk-significant, would you still feel better if you didn't --

MEMBER KRESS: Yeah. I would have still thought the performance was bad.

MEMBER ROSEN: Wait a minute now. Let's use the example you just raised. South Texas has three safety trains, three ESW pumps, call them EC pumps but it's the same thing. So what you do on your bottom line there, your bottom bullet is, you have three, plus two, plus two, not three, plus two, plus one. So you end up with ten to seven, or ten to the minus seven, which isn't white any more. And what it does, it's green. It takes into account the fact that the plant has more redundancy for essential services.

MEMBER KRESS: Yeah, but I would have said that was bad performance in South Texas. It ought to be a bad performance.

MEMBER ROSEN: It was bad performance.

MR. COE: It was. It's a finding.

MEMBER KRESS: Yeah.

MEMBER ROSEN: It's a finding, but it gives the plant some credit for the installed redundancy.

MR. COE: Correct.

MEMBER ROSEN: And you're suggesting we should take that away, and I don't agree.

MEMBER SHACK: It doesn't help its performance.

MEMBER KRESS: It doesn't help the performance. That's right.

MEMBER ROSEN: No one argued that it did. It's just properly -- the redundancy is properly reflected. The plant's owners invested in the additional redundancy. They should get some credit.

MEMBER KRESS: I think the assumption ought to be that poor performance can override a good plant design, and this sort of mixes them up, and I don't think you should mix them up. I think you should have performance being performance.

MEMBER ROSEN: I think the reality of it is you have both performance and design. You can't -- tracks are stubborn things, Tom.

MEMBER KRESS: Yeah.

MEMBER ROSEN: The fact that those pumps are out there, and are installed, and are safety-related, you can't argue them away.

MR. COE: It gives us a more direct link to public health and safety risk, which is really at a high level. What the Commission asked us to do is base our actions more on an objective measure, such as that we could come up with, and this is the one --

MEMBER KRESS: Then we'd fall back on Bob Christie's "Living PRA", and look at the CDF. I think we all --

CO-CHAIRMAN APOSTOLAKIS: We look at the risk.

MR. COE: That's right. We look at --

MEMBER KRESS: Well, look at LERF. I'm sorry.

MEMBER ROSEN: The wind blows.

MR. COE: Well, these are good questions. And actually, on the next page is a list of four of the principal sensitivities that will change these results. And I thought that this was valuable to you, to give you a sense for -- to see how the numbers racked up to give you a white.

If you go down these four bullets, for one, the exposure time was seven days. If it changes

by an order of magnitude, then the result changes by an order of magnitude. Okay? If it was 70 days, you would be talking about a yellow instead of a white. If it was only less than .7 days, we'd be talking about a green. Okay? So that will influence -- the actual facts of the matter will influence the significance here.

MEMBER WALLIS: So if you get Tygon tubing in your pump, it doesn't happen for very long, it doesn't matter.

MR. COE: Not that it doesn't matter. It's just that the significance, if it's .7 days instead of seven days, you would expect the significance to be just under the green/white threshold, which makes it green. It's still a finding. The licensee still has to correct it, but we wouldn't necessarily implement a supplemental inspection procedure. We would allow the licensee's corrective action program to deal with that issue. It's still a finding we still document in our inspection report.

In addition, the mitigation capability you mentioned, if a plant has greater redundancy, that would influence the significance of this outcome. Common cause effect would be an intrinsic aspect of

this analysis. If we found that this Tygon tube was actually -- maybe there were multiple Tygon tubes that impact or that could have impacted all of the pumps simultaneously, that could have been taken into account. There may have been an order of magnitude effect there, and again, it could have bumped a order of magnitude.

Recovery, in this particular case it didn't apply because once the tube was wrapped around the impeller, there was no chance that the operators could recover, so they didn't get any credit for it anyway. Had a different situation arose where there may have been an opportunity to take recovery action, we would have assessed that. And if it was warranted, if we felt it was warranted, we may have given credit, which might have taken that white to a green, if we had given an order of magnitude credit.

MEMBER WALLIS: During this time, and it failed a surveillance flow test, but presumably, the Tygon tubing had been in there for some time before the test was run?

MR. COE: As best I understand it, is that the -- and I'm not sure of the exact time sequence and time line, but somehow they were able to figure out that the Tygon tube fell from its location and entered

the intake structure 172 hours --

MEMBER WALLIS: So they worked back to when it came in.

MR. COE: Yes, sir. Okay. The next example is the tube failure or the tube integrity problems. And although it's not represented in the slides, this is clearly Indian Point Two. In this particular case, and I'm going to make the distinction again. We had an event that initially we thought might be significant because tube ruptures in general are typically -- the events themselves could typically be significant. When we investigated the event itself, we did find some problems with operator response, but on the whole, the actual risk significance of the CCDP for the probability that that event -- given that that event, that the core could have been damaged, was relatively low. However, subsequent investigation identified that the tubes had been degraded over a period of approximately two years. And at the end of that two year period of time, there was a tube rupture event.

This slide here on page 14 identifies that there was a minor radiological release that was within regulatory limits. It was about 146 gallon per minute leak, which isn't -- it's not a double ended single

tube rupture. It was less than that actually, and that has an influence later on here. But there were some identified performance issues and, therefore, the finding though that is the subject of this discussion is the deficient tubes, the fact that tubes were allowed to remain in service over a period of approximately two years in a deficient state. And that that was because of deficiencies involving the licensee's in-service inspection program at their last outage.

The Phase 1 process asks some screening questions again, and in this case the finding contributed to the likelihood of a primary system LOCA initiator, and that automatically requires a Phase 2 evaluation. This is a trigger that we set a low threshold on. A system LOCA is a potentially significant event no matter what the circumstances, and so we want to do further analysis, so we went to Phase 2.

MEMBER WALLIS: So the bad performance was having a deficient inspection program.

MR. COE: That's correct.

MEMBER WALLIS: I mean, this would never have been discovered unless this tube had actually failed?

MR. COE: I don't know that it would never have been discovered, but we would hope that there would be some evidence at some point in time, you know, less than a tube rupture.

MEMBER WALLIS: There could be other plants out there with the same deficient inspection program who haven't yet had a tube failure --

MR. COE: Yes, indeed.

MEMBER WALLIS: You wait until they have a tube failure before you diagnose that they have a red situation?

MR. COE: No. In fact, I mentioned earlier today that the inspection procedure for in-service inspection review has been modified since this event occurred to give added weight and added effort, and further guidance, further detailed guidance to the inspectors, so that we can potentially identify a weak program at an earlier stage.

CO-CHAIRMAN APOSTOLAKIS: So it seems that this event and this occurrence in Davis-Besse have a lot in common. They both have deficient problems, and they both refer to the pressure boundary.

MR. COE: Yes.

CO-CHAIRMAN APOSTOLAKIS: So clearly there's a message there. We have to do something

about it.

MEMBER WALLIS: And Davis-Besse didn't have a rupture.

CO-CHAIRMAN APOSTOLAKIS: No, but the fundamental -- one of the causes was the deficient corrosion inspection program.

MR. COE: Exactly right.

CO-CHAIRMAN APOSTOLAKIS: But the pressure boundary, I think creates a unique problem. I mean, coming back to this earlier discussion and the assumption behind the safety conscious work environment, the corrective action program and so on, that if they are not very good there will be indications, you know, deteriorating equipment and so on. When it comes to the pressure boundary, you may not be able to see that deteriorating until it's too late. It's kind of a unique situation, and we have to pay special attention to it, it seems to me. That assumption doesn't seem to hold very well when it comes to the pressure boundary.

MR. COE: Which assumption? I'm sorry.

CO-CHAIRMAN APOSTOLAKIS: The assumption that I will see deterioration in the performance of equipment if the safety culture is not very good.

MEMBER SHACK: Before something really --

CO-CHAIRMAN APOSTOLAKIS: Before something really bad happens.

MR. COE: I understand. And we have seen evidence before, pressure boundary degradations that have not been -- you know, that resulted in events. Of course, the CRDM nozzle leaking, I think somebody had mentioned earlier, the Surry high-pressure injection nozzle that had the circumferential crack in it. These things cause evidence to occur, high leak rates, high primary leak rates and that sort of thing. And the licensee is responsible to follow those up, and we're watching as they do.

I don't disagree that we perhaps need to be more sensitive to pressure boundary degradation issues, and I'll agree right now that what we do, my hope would be that -- and anybody who can understand that, you know, in a risk model, if you increase the likelihood of a small break or a medium break LOCA, that you get a fairly significant increase in core damage frequency risk. It is fairly sensitive. Core damage frequency is fairly sensitive to those assumptions, and if those assumptions change, if the frequency, or the probability or likelihood of those initiating events increase, then we can easily get to some fairly significant inspection findings. So

knowing that should prompt greater sensitivity to evidence that occurs in a plant that suggests that there might be pressure boundary leakage.

MEMBER WALLIS: How can you relate a deficient program to CDF?

MR. COE: Only through the actual degradation that we know has occurred.

MEMBER WALLIS: You have to then find the degradation. The program being deficient itself has no influence on your CDF, although it may be the root cause of an ultimate problem.

MR. COE: A deficient program raises the likelihood of a greater possibility of an actual impact to plant systems. But unless we find that impact, or identify it, or it self-reveals, you're correct. A deficient program, we can comment that perhaps the licensee isn't following a particular standard, an industry standard, or that they might not be even following their own internal processes and procedures. And those might even be findings, but typically they're not going to be greater than green unless there's been an actual impact on safety function.

CO-CHAIRMAN APOSTOLAKIS: So what was the problem? Why was the inspection program deficient?

MR. COE: In the case of this example.

CO-CHAIRMAN APOSTOLAKIS: Yes.

MR. COE: Partly, I would say because the industry standards for in-service inspection of steam generator tubes is kind of a constantly changing thing that sort of depends on the state-of-the-art. As time goes on, the probes become better, the equipment becomes better, the analysis methods become better. At the same time, plants are different in the way that they apply this equipment, and the way that -- and they analyze the results. And some plants, there may be a lot of noise in the system. There may be -- they were having difficulty discriminating the defects from the noise, that sort of thing. There's a signal-to-noise ratio aspect of this finding that wasn't -- the licensee's noise levels were fairly high in this case.

Again, this is all reflected in the inspection report, and this has been going on for quite a while. But what it has resulted in is additional inspection guidance in this area, and we hope that we're addressing some of these issues, and increasing the sensitivity.

The other thing is, is that not all plants have steam generators that are this old. And all of them that do, are replacing them ultimately, or have

plans to. So, I mean, over time we would hope that the overall risk of steam generator tube ruptures gets better.

In this particular case, the analysis turned out to be red, and in fact, the assumptions that the staff made, you know, were as much related to core damage frequency as they were to large early release frequency.

CO-CHAIRMAN APOSTOLAKIS: Isn't red about ten to the minus four?

MR. COE: The red/yellow threshold for CDF as we know, is ten to the minus fourth per year core damage frequency.

CO-CHAIRMAN APOSTOLAKIS: This is yellow.

MR. COE: No, actually that's -- the large early release frequency thresholds are an order of magnitude lower. And in this case there was a presumption of a one-to-one relationship between core damage frequency and large early release, because if core damage occurred because of a steam generator tube rupture, it would be a direct path to bypass containment through the safety relief valves. And that's a somewhat conservative assumption, perhaps, but it's for simplicity and for, you know, kind of maintaining a standard across at PWRs. It's the way

that we've based our SDP assumptions.

MEMBER ROSEN: Was it influenced by the site population density?

MR. COE: No, sir, it was not. It was only --

MEMBER ROSEN: It would not have been. It would have been red at a site with very low population density, as well?

MR. COE: You're talking about collective risk. And no, sir. The metric is specific to the plant itself, whether there's a large early release potential there or not, or how much of one there is. It does not -- the metric that we've chosen to use does not depend on population density. It's an interesting point, but it -- I'm not sure how we would adjust the -- how we would predictably and consistently make adjustments for population density, because once you start doing that, you may have to take into account prevailing winds and everything.

MEMBER ROSEN: You also have to take into account the definition of LERF. Large early release means before effective -- early means before effective response measures can be implemented. At a site with very low population density, it might have been possible to implement effective response measures, so

you would not have had a large early --

MR. COE: You're exactly right.

CO-CHAIRMAN APOSTOLAKIS: You don't do that on a site-specific basis.

MEMBER ROSEN: I'm just saying if this had happened that --

CO-CHAIRMAN APOSTOLAKIS: Just take the release categories and they say on a generic basis, if this happens --

MEMBER ROSEN: One could argue --

CO-CHAIRMAN APOSTOLAKIS: It's generic. I mean, it's not -- it should be plant-specific.

MR. COE: It may be a future refinement, but right now we did not go to that level or degree.

CO-CHAIRMAN APOSTOLAKIS: So what does red mean now?

MR. COE: For delta LERF it's greater than ten to the minus fifth per year.

CO-CHAIRMAN APOSTOLAKIS: But then the response -- shut them down?

MR. COE: Well, the red in this particular case they shut down to replace their steam generators. But the agency response was an inspection procedure that essentially initiated essentially about a staff year worth of direct inspection effort. And there's

even been some follow-up inspections beyond that that have continued to examine some of their corrective actions and their effort to improve their ISI program.

MEMBER ROSEN: Can I finish my thought about your argument with the large early release? To me, it's the same argument one makes with respect to redundancy. It's a plant feature, the low population density that can't be argued away by semantics. It is, and this goes out and looks at it, so if you take -- if you credit additional redundancy, and getting down to the fine strokes and deciding between yellow and red, for example, in a case like this, one ought to consider the incontrovertible facts of low population density.

MEMBER KRESS: Well, when they looked at the LERF that corresponded to fatality, a safety goal, they found that plants vary about that a factor of four. For LERF they would meet the prompt fatality safety goal depending -- and it's site-specific, but that doesn't really count. That's an individual risk. The LERF is an individual risk, and no matter -- and you're only going to vary a little bit between sites on that because it is an individual risk. One guy there can raise it up, so what they ought to have is something besides LERF dealing with those things, and

not to take into consideration the total population, the total number of deaths.

MR. COE: One death is as bad as two.

MEMBER KRESS: Yes, you're basically right.

CO-CHAIRMAN SIEBER: If you're the one.

MEMBER KRESS: If we're going to stick with LERF, it's all right with me if they want to make it across the board with all the plants. If they want to do something that's more correct, they ought to take into consideration the population.

MEMBER WALLIS: I'm trying to separate out this Phase 1 and Phase 2.

MR. COE: Uh-huh.

MEMBER WALLIS: Their performance, they had lousy performance because they had a poor inspection program. But they could have had a steam generator tube failure in spite of the fact they had an excellent inspection report, that the steam generator tube had just happened. It's no reflection on their performance.

MR. COE: That's correct.

MEMBER WALLIS: And yet on the probabilistic analysis, it still gives the same CDF numbers.

MR. COE: Actually, we would never -- as I mentioned I think earlier, we would never enter the significance determination process unless we'd already determined that there was a performance deficiency.

MEMBER WALLIS: Okay. So it's key that they have this deficient inspection.

MR. COE: Yes, sir, it is. That's the starting point, yes.

MEMBER WALLIS: Although the effect on public safety of having a steam generator tube failure is the same.

MR. COE: Yes, that's correct.

MEMBER WALLIS: So I'm not quite sure how you're balancing risk and performance here.

MR. COE: We're measuring performance using a risk scale. Again, we're forthright and honest. If we have a steam generator tube rupture that's spontaneous and is not linked to any performance deficiency on the part of the licensee, we have programs such as ASP, and we would stand up and acknowledge what the significance, what we felt --

MEMBER WALLIS: But with the green -- you cannot find any --

MR. COE: There would be no finding, there would be no color.

MEMBER WALLIS: You cannot find there's anything they did which led to it.

MR. COE: That's correct, because we're measuring -- we're trying to measure licensee deficient performance, and so you have to start with that assumption. If you talk to the people who have monitored and conducted the accident sequence precursor program, one of the insights that they derived, that they offered at the beginning of the ROP was that that event will happen without any correlation to a plant's performance. That event will happen to good performers with as much frequency as they happen to bad performers.

MEMBER KRESS: In that case, does NRC get a red finding?

MR. COE: That's a good point. And, in fact, if a steam generator tube rupture occurs through no -- because, in fact, the licensee has complied with all regulations and there is no deficiency in performance, maybe the NRC does need to look at the regulations. Maybe the performance levels and the standards and requirements should be tightened.

CO-CHAIRMAN APOSTOLAKIS: This is the classic question in quality controlling. Something extraordinary is observed. The fundamental question

is, is it due to a systematic cause, or is it random events, that you have to make a judgment as to what it is. That's what you guys --

MEMBER ROSEN: The way this is said is if you flip a coin ten times and it comes up heads ten times, you have witnessed a rare event.

CO-CHAIRMAN APOSTOLAKIS: Or is the coin biased. That's a question. Is it biased, or have you witnessed a rare event? Do you think that all these problems with the pressure boundary would go away if the material experts did a better job?

MEMBER KRESS: Are you being Dana Powers now?

CO-CHAIRMAN APOSTOLAKIS: I'm asking Doug for an answer.

MR. COE: If he materials -- what, the materials organizations in NRC, or the licensee materials, the vendors?

MEMBER ROSEN: It's intended to provoke our materials expert.

CO-CHAIRMAN APOSTOLAKIS: I managed to provoke one. The other one --

MR. COE: We can always improve.

MEMBER SHACK: He works on BWRs. That's his solution to the problem.

MR. COE: What is needed on the part of both the industry and the NRC, is an aggressive effort to find out the causes, and to understand the physics of failure when these things occur. Every failure provides a window of opportunity to increase our understanding. And if we don't take advantage of those windows of opportunity and really seek to understand the physics of the failure, then we can't decide whether our programs are good enough.

Let me move on to the next example.

MR. FRAHM: Did you want to go through this?

MR. COE: I think we did. We already covered -- all of those sensitivities apply to all of these reactor safety examples, and can influence them. They are the principal means of influencing, and I offered them to give you a sense of sensitivity, things that can change these results.

Example C, starting on page 16, was a loss of instrument air, but in fact, this is also turns out to be a red issue. And again, although we haven't indicated it here, it's clearly the Point Beach. In this case, the loss of instrument air actually has an auxiliary feedwater system because the minimum flow recirculation valves all fail shut on loss of

instrument air. And if the pumps are being utilized to restore and maintain steam generator level, and the recirculation valves shut, at some point the operators throttle back on the flow to the steam generators, and then there's no -- and if there's no recirculation flow, the pumps will burn up within a very few minutes.

A number of things may cause a loss of instrument air, in addition to a spontaneous loss of instrument air, and that could be caused also by a loss of outside power, loss of service water, or a seismic event. These were considered during the SDP. This condition was present since the initial start-up, so in such a case we annualized the annual risk on a per year basis. We don't try to accumulate risk over prior years. Essentially --

CO-CHAIRMAN APOSTOLAKIS: The crucial step I thought was always is th is a performance issue.

MR. COE: Correct.

CO-CHAIRMAN APOSTOLAKIS: Why is it a performance issue?

MR. COE: Well, that's a good question. I guess I could look up the specifics in here, but I'm going to speculate just a little bit that -- I don't know how it's actually articulated in the official

documentation, but this is a design deficiency in which there was a number of opportunities over the period of the plant's operation since start-up to identify this. It's essentially a failure modes and effects analysis kind of a result, where you conclude that there's a -- to be a substantial impact, risk impact or safety impact due to the single failure.

MEMBER ROSEN: This is a license design. Right?

MR. COE: Yes, this is a license design.

CO-CHAIRMAN APOSTOLAKIS: He's not saying that the performance issue was the design itself. It's the failure to find the deficiency --

MR. COE: That's correct. And I believe that's the way it's articulated. In fact, the license --

MEMBER LEITCH: Also, with respect to the lack of the operating procedures warning the operator about this potential. I thought it related -- the way it got to performance was through inadequate operating procedures.

MR. COE: It could. That may be. It actually was identified by the licensee's PRA staff, by the way, but it was -- the conclusion I think that the staff drew was, that they had a number of

opportunities up to that point.

Now the question here as to whether or not this can -- this was -- there is a provision, and I'm not really prepared to talk about it here, that this was a -- could be considered an old design issue. There's some credit that can be given under the terms of our assessment process that allows some consideration of the fact that they found this through a program, or through a means that was over and above the normal routine expectation that the agency has for these kinds of activities, design review activities. That decision hasn't been made yet. Okay.

Whether we -- and there's -- I'm not prepared to go in all the reasons why, because that's still pre-decisional, but there is a finding here, and it does relate to missed opportunities to identify this condition. I can't put my finger on it in the package right here, but --

CO-CHAIRMAN SIEBER: It seems that as we go along through the process, one of the deficiencies of reactor oversight, whether it's this program or the SALT program, or anything else, this one in particular is that it's not particularly timely. You know, the event occurs or the deficiency is found, or a violation is found, and if it has more than one order

of complexity to it, it seems to take forever. You know, it seems to me that that's not good. It's like spanking your dog two days after he wets on the carpet. And maybe -- is there some hope that the process would ever speed up?

MR. COE: Yes. The answer is yes. And we've acknowledged from an early point that we need to improve timeliness. The Commission has reminded us of that. The implementation of the SDP improvement initiatives are designed to deal and address each of the elements that we see as providing untimeliness, a factor of untimeliness. Part of it involves just getting more clear on what the risk characterization process is or should be. And coming to perhaps a better balance of how detailed our analytical calculations have to be relative to the judgments that are being made, and all of the uncertainty that exists, that we acknowledge exists, both the epistemic and the aleatory, and to be able to continue to get to a decision point even in the face of those uncertainties. As long as we recognize them, we acknowledge them, and we agree that we can make a judgment and move forward.

Now it is always the staff's judgment. We invite perspectives from the licensee because they

often have good information to provide, and so we do solicit and invite that. Our program, you know, allow for that, and in fact, requires it. So can you get to a more timely result? We're going to try to find all the things that we can do to improve the efficiency to get to a decision faster.

CO-CHAIRMAN APOSTOLAKIS: So you did the Phase 2 and you concluded it was red. I think the message here is that we are focusing on the fact that there was a performance issue, because they missed a number of opportunities for finding those design deficiencies. But at the same time, we're saying look, this is not like the old SALT or other ways we used to use, where just the fact that they missed it is good enough for taking some action. The fact that they missed them, and it was a safety-related issue makes it important, so in that sense the process is focusing on performance, but is risk-informed. That's the way I see it.

In other words, the calculation of the red only sends the message that for certain things you have to be more careful than others. Just like missing things may be, you know, you missed something but it was not important. That's fine. This is an industrial facility, after all, but when it comes to

safety, you know, you have to be risk-informed regarding what you're missing.

MR. COE: I would agree, except I would say it's not that it's fine. They still --

CO-CHAIRMAN APOSTOLAKIS: Oh, it's not fine, but it's not of the same importance.

MR. COE: It's not of the same importance. And if we act as an agency in a risk-informed fashion, then there's an expectation, a natural one that the licensee will also act in a risk-informed fashion, will pay more attention to the things that are more important.

CO-CHAIRMAN APOSTOLAKIS: But my point is that the focus here should not be on the red. The focus should be on the original cause that you identified, which occurred in a circumstance that was risk -- I don't know if it's significant but relevant, risk relevant. If you put it that way then I think you're really focusing on performance throughout. And risk is just a supplementary piece of information that helps you discriminate as opposed to the old case where a violation was a violation. Missing something was missing something, independently of its significance.

See the danger that I see here is because

of these equations, and the two, and the three, and the parentheses, and blah, blah, blah. Maybe people will focus too much on this stuff, forgetting the reason why we're doing all this.

MR. COE: That's a good point, and the focus needs to quickly get to an assessment of how we grade the significance of this issue so we can move, so the licensee can move on, we can all move on to correct the problems. Okay? Because that's our ultimate intent, is that the licensee correct these problems. And so I don't think I would disagree with anything you say. I think that's what we're trying to achieve. If we act in a risk-informed fashion, the licensee will act in a risk-informed fashion too. That's our goal. And so I would have to agree.

CO-CHAIRMAN SIEBER: What I'd like to do is, being that lunch time is fast approaching and we have a number of examples to go, it would be good if you could finish up instrument air, and perhaps do one other.

MR. COE: Sure.

CO-CHAIRMAN SIEBER: And the one that I would be interested is Example F.

MR. COE: F.

MR. FRAHM: Actually, that's going into

this afternoon's portion.

MR. COE: Is after lunch.

CO-CHAIRMAN SIEBER: Oh, okay.

MR. COE: It's after lunch. We'll get to that.

MR. FRAHM: Doug is only handling the reactor safety SDPs which include the first five, so there would only be two additional ones.

CO-CHAIRMAN SIEBER: WE'll deal with that --

MR. COE: I only have -- I'm at the end of this one.

CO-CHAIRMAN SIEBER: Okay.

MR. COE: And I just have two more, and they're relatively simple, I think.

MR. FRAHM: It will be right after lunch.

MR. COE: The loss of instrument air is represented here in a Phase 2 level of detail just to give you a sense of where the -- what the significance, the risk significance derives from. And in this case, the accident sequence of greatest concern is the loss of instrument air, the spontaneous loss of instrument air, and with no remaining aux feedwater capability. That was confirmed by the licensee's more detailed analysis, using the more

detailed risk models, and our own, as well. But this is a high level representative of the drivers, the risk drivers for that issue. And that's really all I need to say about that example.

The next example is a little bit of a different one. It's captured under the mitigation cornerstone because it's operator requalification or operator performance kind of a deficiency, and operators in this context are considered part of the mitigating strategy or mitigating systems of the plant.

In this particular case, the SDP was developed in consonance with some industry dialogue. This was -- there was an opportunity for the industry to comment and interact with us as we developed this particular SDP, and it's fairly cut and dry. And essentially --

CO-CHAIRMAN APOSTOLAKIS: Could you send it to us? I'm curious how you developed the risk metric that reflected this particular failure. What did you do, you changed the operator error rate?

MR. COE: I'm not -- no, I don't know. Let me put it that way. I'm pretty sure that we did not change the operator fail rates because that is not part of the basis, I think, that we provided in the

basis document, although it's been a long time since I read that portion of the basis document.

I don't know to what extent you've had a chance to examine this particular SDP, other than --

CO-CHAIRMAN APOSTOLAKIS: Can we have this SDP? Can we have it sometime in the next couple of weeks?

MR. COE: Yes, absolutely.

MR. FRAHM: 609, Appendix I.

MR. COE: Appendix I.

CO-CHAIRMAN APOSTOLAKIS: What does IMC stand for?

MR. FRAHM: Inspection Manual Chapter.

CO-CHAIRMAN APOSTOLAKIS: No we can look at the chapter, but I would like to see the actual SDP.

MR. COE: It's Appendix I of Manual Chapter 0609.

CO-CHAIRMAN APOSTOLAKIS: It's based on Appendix I, but can I see the actual SDP for this event?

MR. COE: Actually, if you turn the page to the next page, there's a table which essentially represents the SDP. The particular issue in question was the high failure rate during annual simulator

examinations as part of the licensee requal.

CO-CHAIRMAN APOSTOLAKIS: I understand that, but SDP produces CDFs.

MR. COE: Not in this case. This is an example of essentially of a performance-based SDP in which there really wasn't a good mechanism across the board to create a generic SDP -- I'm sorry, to create a plant-specific SDP for these kinds of issues, so a generic SDP was created, and it was built from essentially judgment, and not from a particular risk analysis or evaluation.

CO-CHAIRMAN APOSTOLAKIS: But you have observed a high crew failure rate.

MR. COE: Yes.

CO-CHAIRMAN APOSTOLAKIS: Why did you need a color to decide. Why go through the pain of developing the color, since it's something that's really very difficult to quantify. Did you gain any additional insights or did you decide your first reaction was to do AB, and then the color says oh, no, you should also do C and D? I mean, in a pragmatic way again, do we always have to develop a color?

MR. COE: When we have an inspection procedure that goes to look at a licensee activity that's governed by our regulations, there should be a

way of adjudicating the findings that come from that in terms of their significance. And I will admit that we don't have that in all cases. We don't have necessarily an SDP for spent fuel issues, for example. So we're continuing to work on those kinds of things, but in this particular case we do inspection of requalification programs, and we generate findings. And in this case, the operator licensing people who manage this program felt that they needed -- that this was an SDP that they needed in order to adjudicate the findings coming from that inspection. And when we find high failure rates, it certainly prompts our questioning and our evaluation, and so we needed a consistent predictable scrutable way in which we can grade licensee performance. So we account for in this SDP, if you'll notice on the table, we account for the fact that licensee may have any number of operating crews, and so we gauge our significance characterization on the number of crews that failed our simulator exam relative to the number of crews that they have, so it's like a percentage.

MR. SATORIUS: Doug, if could help here too, our operator licensing person is not here. Apparently they went to lunch, but in the past, we would perform examinations of requal programs. And

occasionally there would be unsat requal programs, and there was a certain level of effort of follow-up inspection that was performed as a result of those unsatisfactory requal programs.

My thought is, is that this table captures what had been learned through experience of examining requalification programs, determining if they're satisfactory or not, and what levels determined when they were unsat, we would undergo a certain inspection effort to assure that they reached the quality that would be considered satisfactory again, so that's what this table was derived from, that experience that was gathered through inspecting requal programs.

CO-CHAIRMAN APOSTOLAKIS: How come there is no red? And you guys are resisting so much removing the reds from the performance indicators.

MR. COE: Some performance indicators don't have red values either. This was a case where the level of effort --

CO-CHAIRMAN APOSTOLAKIS: Well, we have a precedent. Now we're negotiating the price. Can you remove it also from the frequency of initiating events since you've already done it?

MR. COE: This particular SDP, I think that the judgment was made that the 95-003 level of

effort, which again constitutes about a staff year worth of direct inspection effort, not to mention all of the documentation and prep that goes with that, was too much. It wasn't necessary to focus on a very specific program that had fairly definite boundaries.

CO-CHAIRMAN APOSTOLAKIS: Twenty-five transients is too much. It's the same logic.

MR. FRAHM: Point taken.

CO-CHAIRMAN APOSTOLAKIS: It's the same logic.

MEMBER KRESS: On this table you here, this matrix, give me a little bit of information on the vertical axis. For example, if I look at the four or five level on that vertical axis, does that mean that plant only has five operating crews, or does it mean that they only gave five tests to the number of operating crews they had?

MR. COE: The answer to that should be in the definitions for this SDP, and all I'm showing here is the table, so I'm at risk of giving you the wrong answer if I try to --

MR. SATORIUS: I can help here, and that is the requal -- the regulations that require operators to undergo a requal program, and I don't know that periodicity, but they don't have to do it

every year, so that would --

MEMBER KRESS: But there is a control over -- given how many crews they have, there's a regulatory control over how often they have to be tested.

MR. SATORIUS: That's correct.

MEMBER KRESS: So you don't have to --

MR. SATORIUS: So in other words, a facility may have, I'll just pick a number, 13 crews. And once again, these are just illustrative examples. Five or six may have to every year cycle through a requal program, so that's what you get for the left hand. That's the number of crews that took the test.

MEMBER KRESS: And I would have thought that might be a performance indicator as to whether they actually did that, but I presume there's such controls on that that there's no way they'd miss --

MR. SATORIUS: Well, I wouldn't say no way because I was involved on July the 4th on an issue at Dresden where we had to issue 53 notices of enforcement discretion because the licensee had read the dates wrong and failed to administer requal exam within the periodicity.

MEMBER KRESS: Now that to me would have been a performance indicator.

CO-CHAIRMAN SIEBER: Just to clear up the record, every operator who is licensed goes to a requal program every year.

MR. SATORIUS: That's true.

CO-CHAIRMAN SIEBER: And it's a licensee run program. And there is an exam associated with that program, a simulator exam and other exams. And on the other hand, the NRC oversees a certain portion of those every year, and so this comes to the portion that the NRC oversees.

MEMBER KRESS: They also --

MR. SATORIUS: That's a good clarification.

MEMBER KRESS: They also approved the licensee's specific tests, don't they, before?

CO-CHAIRMAN SIEBER: That's correct. You submit and they say yes or no to the questions.

MEMBER WALLIS: I'm very surprised at the levels here as a naive member of the public. If part of them fail you give the green. If a third of the school bus drivers fail their driving test in my town, I don't think that's an insignificant event.

CO-CHAIRMAN SIEBER: Don't ride the bus.

MEMBER WALLIS: Why are you so soft?

MR. COE: Actually, the particular plant

in this example didn't think -- thought we were pretty harsh in awarding a - what was it, a yellow?

MEMBER WALLIS: I would think if one of them fails, it's a significant event.

CO-CHAIRMAN SIEBER: Well, what happens is that the operator who fails cannot operate until he undergoes remedial training and takes another exam. It's like the school bus driver who just got his license revoked - okay - or suspended until such time as he could demonstrate or she can demonstrate that they can operate --

MEMBER WALLIS: This guy has been operating until he took the test.

MEMBER LEITCH: That's right. What this is, is number of crews too. This is not particular operators. I mean, we're talking here about simulator performance, so what you do is evaluate the crew competence, not particularly an individual -- not necessarily -- in fact, not at all an individual operator. You're looking at the performance of the crew on the simulator which may be a licensed operator and an STA or something in the simulator.

MEMBER WALLIS: Well, I don't know what the test is, but if it means that if they were faced with an accident that 30 percent of the time they'd

make the wrong decision and you still give them the green, that doesn't sound good to me at all.

MR. COE: And actually, in this particular case the licensee thought we were harsh because the reason that they failed their operators in these cases weren't necessarily because they failed to perform critical tasks correctly. There were infractions of lesser significance that they used in their own evaluation process to cause them to fail. And so part of that argument coming back to us was that well, you know, they really didn't fail anything really critical, and we -- you know, we set a higher standard for ourselves, so they thought they'd actually get some credit for that. But we established the SDP based upon their own determinations of their failure criteria.

CO-CHAIRMAN SIEBER: I would point out that we only get 30 minutes for lunch today, and if we break right now we'll just get the 30 minutes. Any further discussion beyond this will encroach on that length of time. Now I don't think there is time to talk about fire suppression. We have an hour after lunch. You amongst yourselves of the staff can decide whether you can deal with EP, rad con and fire suppression at the same time.

MR. COE: At the end of the day --

MS. WESTON: You have one hour after lunch, and then one hour after our break, so you know we have two hours to finish your's.

MR. COE: And if at the end of that time you want to come back and look at this example, we can do that.

MEMBER ROSEN: Given the fire protection subcommittee's comments on fire suppression and SDP I would particularly like to go through this one.

CO-CHAIRMAN SIEBER: Well, why don't we take our luncheon break now and come back at 1:00, and then we could continue on where we're at.

MEMBER SHACK: Will we release Doug if we go through this one now?

MR. COE: No, I'll come back.

MEMBER SHACK: You'll be back in.

MR. COE: Yes, I'll be back.

MR. FRAHM: Doug is a key member of the team.

MR. COE: I'll be happy to cover that example.

CO-CHAIRMAN SIEBER: Okie-doke. Okay. Let's recess until 1:00.

(Off the record 12:32 - 1:07 p.m.)

CO-CHAIRMAN SIEBER: I think we have our discussions on the ROP. And we'll start with fire suppression since everybody seems to like fire suppression.

MEMBER ROSEN: Better than fire going out of control.

CO-CHAIRMAN SIEBER: Well, it depends on the fire.

MR. FRAHM: And in the interest of time,over the next hour we hope to cover this example, as well as examples in occupational and public radiation safety, so we definitely need to keep moving.

MR. COE: I'll just preface the beginning of this example by saying that as you probably know, the fire protection SDP continues to be under intense review to seek ways in which it can be improved in terms of its efficiency of use and simplicity, and its overall usefulness and effectiveness. That work is ongoing.

The example here is a reflection of the existing process as it's currently documented in Manual Chapters 0609, Appendix F. The deficiency in this particular case was the revelation that a particular fire area which housed a number of

components used or credited for safe shutdown did not have the required suppression equipment installed. In fact, I believe the licensee in conducting some follow-up research to a tri-annual NRC fire protection inspection determined, made the determination that this fire area has not been correctly classified, and therefore, did not have the correct suppression equipment, so they placed the issue in their corrective action program. But subsequently, they closed out the issue inappropriately before they had addressed the need for the additional suppression equipment. And it was reopened after the NRC identified the inappropriate closure in a PI&R inspection. So here's an example, I think, that reflects our earlier discussion this morning in a case where the NRC identified a closed issue that was closed inappropriately, and subsequently the licensee reopened it. That finding was made through the PI&R inspection procedure.

In this particular case, the equipment that was in this room included the B train motor-driven AFW pump, the turbine-driven AFW pump, two 480 volt switch gear buses and an instrument air compressor. And there were cables for both A and B trains of equipment that passed through this fire

area. It's kind of hard to imagine that they would have missed that.

MEMBER ROSEN: And they shot themselves in the foot. They might have had an old design issue if they hadn't then shot themselves in the brain with not correcting it.

CO-CHAIRMAN SIEBER: There were many designs from the 1960s/early 70s that were like that, unfortunately.

MR. COE: I believe that this was an older vintage plant. In any case, the finding then was one of not having provided appropriate fire suppression capability, and that this was seen as a performance deficiency. It entered the Phase 1 screening and passes directly to Appendix F, which deals with findings involving degraded fire suppression barriers and equipment.

Appendix F then goes through some further screening, and it took the issue to a point in the Phase 2 analysis that required some risk evaluation. And that was based principally on a couple of important assumptions. One was the ignition frequency for that fire area, and although it's not given in this slide, I only have a very high summary here, high level summary, the ignition frequency was based on a

value that the licensee used in their own evaluation of this issue. And they got that frequency from an EPRI database that reflected turbine building pump fires. And so that was approximately one to the minus four.

And then there was some credit given from annual suppression, and as it's noted here on the slide, but no credit for any fire barriers or automatic suppression since, of course, they didn't exist. So with an additional ten to the minus one essentially credit for manual suppression, the initiation frequency multiplied the manual suppression gives you an order of magnitude of about ten to the minus fifth.

Then one more factor is involved here, and that is, the ability of the operators to recover one failed train, so if a fire occurred there was apparently in this particular instance an opportunity for the operators to recover one failed train of alternative safe shutdown, and so an additional ten to the minus one credit was given for that recovery. This all, by the way, is in accordance with the prescribed amounts of credit that are defined in this SDP.

Given that, the range of the -- or I

should say the value of the finding in terms of risk significance comes out to be on the order of between ten to the minus fifth to ten to the minus sixth, which is white.

The licensee's own analysis using more detailed techniques involving severity factors and so forth came out to within the same range, at the high end of the white, but still within the white range. So in this case, the Phase 2 result did comport with the licensee's own evaluation, using more detailed analytical techniques.

MEMBER WALLIS: When you say high end of the white, do you mean it was almost yellow?

MR. COE: The licensee came out around ninety to the minus six. Phase 2 doesn't make distinctions any more refined than orders of magnitude.

MEMBER ROSEN: It was still white.

MR. COE: Yes, sir, still. It was an agreement.

MEMBER ROSEN: I was at the fire protection forum, the last one where they showed -- one licensee showed how they had done some detailed fire model given the circumstance. I don't know whether it was this one or another one. I mean, I

know what they were modeling, but I don't know what this one was and so I don't -- and I don't want to. But the question really was about the detailed fire model. Would you have been willing to entertain the discussion of a detailed fire model of this if the licensee had chosen to provide one?

MR. COE: Yes.

MEMBER ROSEN: What would you have done with a good detailed fire model?

MR. COE: Well, I would suspect that in this case a detailed fire model would get to questions of, you know, is there sufficient combustible material in this, or initiators, fire initiators in this particular fire area. And in this case, there were some documented assumptions. I didn't mention it, but regarding that there was sufficient combustible material and sources of ignition that there was a reasonable fire scenario that could evolve to impact the equipment in that fire area.

The kind of modeling that I think you're speaking of, and we've had these discussions with our fire protection staff, you know, often involve the quantities of combustible materials and the location of those sources of combustible materials and sources of ignition relative to the various equipment that

could be impacted, so because of the spatial arrangements it could become very complicated. But it involves, you know, not only the opportunity to combust this material, but also the development of hot gas layers that rise to the ceilings and impact cable trays and that sort of thing. So fire sciences is clearly a complex area. I believe based on the little bit that I've seen that it's tantamount to the severe accident phenomenology that we deal with in terms of its, you know, the various physical -- the physics of what's actually -- what we're trying to model and what's actually happening, and so it's a very difficult area.

We use the best insights that we can to construct this SDP in a manner which lends some structure to our decision process, and that's where we're at.

MEMBER ROSEN: Well, I think that's a good answer, but I wouldn't agree that it's the same as severe core damage phenomenology, because in that case you don't -- you have almost no testing and no experience. And here we have fire, we have lots of testing, and lots of experience with hot gas layer propagating, and testing can be done at reasonable costs and that sort of thing, so there are some real

differences, Doug. But one of the things you can do with a detailed fire model is get some insight into how long it takes for the fire to progress to where more than one train of safety equipment is damaged, and the likelihood that manual suppression, there was no fixed suppression installed, the likelihood that manual suppression could be employed in the time available based on the fire model.

In this case where you've given credit for manual suppression already, I don't think that helps so, you know, this seems like a case where a detailed fire model would not have helped.

MR. COE: And we picked this case because it was relatively simple. Other cases do become more complex and may depend more on the factors that you've mentioned, so your point is a good one.

MEMBER ROSEN: And your answering that if given certain circumstances, and faced with a yellow or some other color finding that the licensee did not want to have and didn't believe was appropriate, because he could have put that fire out, this postulated fire which, of course, is all it is. He could have put that postulated fire out he thinks, and he's willing to do the work to show you a good analysis that under those circumstances he would take

it into account.

MR. COE: And in this case, I think that we certainly -- we did credit the manual actions based on whatever inputs they gave us and our own judgment that the manual actions could reasonably be accomplished so you're right. And we've engaged licensees, particularly in fire protection areas, in which they've expended a great deal of effort to provide to us the results of various tests and modeling, and so forth. And this is causing a lot of concern because of the expense that's required to answer some of these fire science questions, as well as some of the probabilistic questions. So one of the efforts -- one of the objectives of the effort going on now to improve the SDP in this area is to help improve the timeliness and the efficiency of doing these SDPs.

MEMBER ROSEN: Well, I would applaud that, of course. But I also would suggest that if the staff takes a positive attitude towards fire modeling, that the industry is more likely to do it. And doing it reveals a lot of useful things about how fire propagate, both for design purposes and for suppression and operational purposes. And I think the agency ought to encourage that, rather than take a

stance that discourages it.

MR. COE: I agree completely. I think, in fact, I would hold up the difficulties we've had with fire SDP as a really good illustration of why it's necessary to have the engineering and science, fire science people interacting very closely with the probabilistic risk people. In many cases, at least at the initial outset, it seemed like there was a difficulty in communicating across this barrier. But as both sides contributed to the discussion and the dialogue, what's come about today is a very integrating working group of people from both sides of the fence that are working together to try to create and SDP process, and improved SDP process in this area that accommodates the fire science views, as well as the probabilistic framework, so it's a difficult process but it's necessary when we're dealing with this kind of analytical tool. And that's all I have for this example, unless there's other questions.

MR. FRAHM: Okay. Next we have Roger Pedersen to go over some occupational radiation safety issues.

MR. PEDERSEN: Yeah. My name is Roger Pedersen. I'm the subject matter expert in the occupational radiation safety cornerstone to ROP.

Before I start into the specific example which I believe is Example 4 in the package. It says Example F in the slides, but before I go into that, I think I need to talk a little bit about the basis for the SDP in ALARA before we actually get into the example.

A number of the discussions that I heard this morning were reminiscent of a lot of the difficulties that the staff had early on in this process when we were trying to develop the ROP, both performance indicators and the significance determination process. As a matter of fact, the `98 white paper that the industry provided prior to the original public workshop that kicked off the development of ROP, specifically excluded radiation protection, both occupational and public, and security and safeguards from this ROP process, because they were using the definition of risk-informed that was using risk insights from a PRA. And, of course, it doesn't apply to our areas.

The NRC took a broader definition of risk-informed, and that's one that takes risk insights from other sources other than PRA, and we were all excluded. That's why we have a separate cornerstone process.

What that did is force us subject matter

experts and the industry into trying to evaluate how risk is associated or is reflected in our regulatory and licensing requirements. In terms of radiation protection, our measure of risk is dose, so our SDP is somewhat dose-based.

Now from the outset, I'll tell you that we -- there was never any attempt to try to normalize between the cornerstones. In fact, even within our cornerstone between ALARA, which the metric is actually collective dose as opposed to an individual exposure situation where the dose of the individual is the risk determiner. There was no attempt to try to normalize those.

The way we came to the decision gates in the SDP and it was also reflected in how we picked the criteria for the performance indicators, was driven by the action matrix. The action matrix was already developed. There were bins of NRC performance, or excuse me, NRC response that were already pre-identified, and that we went through several public workshops and public meetings to come up with an expert opinion, if you will, subject matter expert opinion as to what level of dose, what level of a performance deficiency that had a certain dose consequence or potential dose consequence to determine

what bin or what NRC response mode we should be in. So having said that --

MEMBER KRESS: We don't think you need to apologize because we think --

MR. PEDERSEN: No, I'm not apologizing. I'm just saying --

MEMBER KRESS: WE think that's the way it ought to be.

MR. PEDERSEN: I'm not apologizing.

MEMBER ROSEN: You need not apologize for some of this.

MR. PEDERSEN: I'm not apologizing. I'm just going through how we came to where we are, and why the SDP looks the way it does.

MEMBER ROSEN: Now you're just redefining risk as not core damage risk.

MR. PEDERSEN: I don't know if that's redefining it or not. The dose limits that we have in Part 20 are based on epidemiology. They're based on mortality and morbidity probabilities of certain dose levels. It's not determined through PRA, it's determined through epidemiology.

MEMBER ROSEN: But the word "risk" and risk-informed regulation has always been meant by those who speak it and those who hear it, to think of

core damages.

MR. PEDERSEN: That's right. That's exactly right.

MEMBER ROSEN: And you're saying well, yeah, but there's another kind of risk. There's individual risk --

MR. PEDERSEN: That was told to me.

MEMBER ROSEN: That's okay.

MR. PEDERSEN: And that's the ground rules that we operated under.

MEMBER ROSEN: Okay.

MR. FRAHM: And we did convey that in our December 19th paper also.

MEMBER ROSEN: An okay kind of thing to do.

MR. PEDERSEN: Okay. ALARA has a very particular place in ROP. It's an exception to just about everything ROP stands for, I think, in that the regulatory requirement to begin with is performance-based. It's a program base. We have a regulatory requirement that a licensee have a program to demonstrate or to provide doses through ALARA, not that the doses themselves are the minimum possible achievable. That's in the Statements of Consideration in the 1994 rule making that established, if you will,

the ALARA requirement, so we had some difficulty.

The industry early on recognized the subjective nature to ALARA. There was no performance indicator that was put forward, and there is no performance indicator in this area. It was left to the inspection program to do the assessment of this area of the radiation protection program. That "admittedly subjective criteria" that's on the slide, that comes right out of the Statements of Consideration in the 1994 that's referenced in the Federal Register right above it.

So we had, as I said, many stakeholder meetings in which we wrestled with how we were going to come up with objective criteria to judge or assess the performance of a subjective area. A number of issues we had to deal with was what is the unit of performance that we're talking about. We're talking about a rolling three year collective dose which was a performance indicator that was previously used in the industry, or are we talking about the performance at any particular outage, or any particular annual cycle?

What we ended up with was -- well, and then a standard to judge that performance against. What we ended up with was coming up with the standard

of the licensee's own program. We judged the licensee's performance against their own program, against the planning that they put into place prior to going into the work activities. They're required -- this is the requirement in the regulation to have a program to determine what the doses are going to be, and if necessary, take actions to minimize those doses, or to reduce those doses, so the outcome of that planning program is what we used as the standard to judge the performance of the licensee's program. And we determined that that was best suited, since the SDP process is supposed to be putting risk-significance to inspection findings, that that would be judged on a planning unit basis.

Early on we used the term "job", which became a major stumbling point in the Callaway enforcement action. There are different definitions of what a job is. The term "job" refers to different things, especially in outage planning. You have a JCN sometimes, that talks about jobs as far as critical path flow and that type of thing. The job that we were referring to here, and subsequently have changed the terminology to a work activity, that's the job or the unit of work that the licensee themselves has broken their outage into for the purposes of ALARA

planning. So we judge on a work activity basis the licensee's performance.

Another issue that we had to deal with was, in fact, that the overall industry performance in ALARA has actually been getting better and better over the last 15, 20 years. We did not want to all of a sudden start trying to put an oar in the water and drive anybody's program, because the overall performance is very good at this time.

When I first got to the NRC back in the early 80s it wasn't uncommon for BWRs, in particular, to have 1,100, 1,200 person-rem outages. The end of the 90s, 1999, Quad Cities had a 600 rem outage and they were very shocked by that. They were embarrassed by it, and I heard the RPM give a presentation at the HP Society Meeting, and there was a ripple that went through the audience actually, because a 600 person-rem outage was now unheard of.

So what we tried to do is provide a process in which licensee performance not only was judged against their own planning, but against the industry, it says "industry average". We actually used the median values in 1999 for the data that was available for the rolling three year averages. The 135 person-rem for a BWR and a 240 person-rem for a

PWR. Excuse me, vice versa. 135 for a PWR and 240 for a BWR. That was the median rolling three year average collective dose for those two classes of licensees in 1999, so the data we had at the time was 1998, was `95, `96, `97.

MEMBER ROSEN: In the case of PWRs, that includes two different basic groups, ones that have extensive steam generator work and ones that don't.

MR. PEDERSEN: We couldn't define it that finely. The data we had was only stratified on BW and PWR.

MEMBER ROSEN: Well, I know you had the data, but I'm saying that you really have two -- because the steam generator work is typically the highest dose activity in an outage, plants that have recently replaced their steam generators who don't have a lot of work to do end up with low levels of rems.

MR. PEDERSEN: This is a very roughing filter, if you will. As a matter of fact, it becomes a filter. What is being shown here on this slide is the original, it's called Group 2 Screening. That grouping is not important. It's part of the -- it's how it's characterized in the Manual Chapter.

Originally, the very first draft of the

ALARA SDP, these two boxes were actually in the SDP as a screening process. The first box is where we're judging the licensee's program against itself. We're comparing the actual dose, collective dose that was experienced for work activity against what was planned for that work activity, and the criteria 50 percent is just expert opinion. Then we go to that second box which is how they stand against the entire industry in terms of a rolling three-year average collective dose.

Historically, that has been -- that rolling three-year average has been a performance indicator. One of the things that the industry stakeholders pointed out in this whole process is that it's been misused quite a bit. That rolling three-year average has a lot of detail in it that is completely covered up by averaging these three years in terms of what a challenge is, whether you have steam generators to replace, or whatever the issue is. And it came up again when we were having stakeholder meetings post the Callaway.

The industry objected to the staff's characterization of people that are -- licensees that have experience, a rolling three-year average above that median as having a bad or a poor performance. And what we determined is actually this rolling three-

year average is more an indicator of the challenge of the program, of the licensee's program, as opposed to the performance of the program. Licensees with a high rolling three-year after collective dose may, in fact, have the best program in the country, but they might have a legacy problem. They might have a problem with poor fuel from early in operations, or whatever the issue is, so it still works out the same.

What we're doing is those licensees that have less of a challenge, that are below the median value that's listed there, the max now at this time we screened them out as having no finding at all, is one of the things we changed in the lessons learned from Callaway. Now it's incorporated in the SDP that's in your package, that indicates that they could have a maximum of a green finding.

The last diamond at the bottom there is just a lower discriminator. We didn't want to be nitpicking the licensee's programs, so the work package, the actual dose that's experienced from a work activity has to be greater than 5 person-rem, if you will.

Now we didn't try to use any risk factors to those person-rem to come up with some absolute risk. This is all expert subject matter -- subject

matter expert opinion as to what levels of issues should make it into the SDP. And then in the SDP, what levels of issues should cross from a green to a white performance issue.

MEMBER LEITCH: That first diamond has the potential to have unintended consequences with a high estimated dose to begin with.

MR. PEDERSEN: And that's one of the things we had to clarify. It's basically -- the guidance given to the inspector is to use the licensee's program outcome, but he needs to review the bases for that, and if he sees a discrepancy in the licensee's historical dose for that job and this planning, he needs to investigate that. And if there is no bases for that, if there is some padding, if you will, of the dose, then he's to use the historically justified dose for that job to base it against.

MEMBER LEITCH: Does this all factor in the -- there's an economic trade-off for man-rem saving. I forgot what the number is, $10,000 of man-rem or something like that is a number that's --

MR. PEDERSEN: It was originally $1,000. We put out a new reg that says $2,000. Licensees use anywhere from 10 to 25,000 dollars per man-rem. That should be factored into their ALARA planning. And the

fact that there isn't a single number also is one of the reasons why we are using the licensee's own planning process as a standard to judge their performance against. We don't know if 25 rem for this job is the right number or not, taking into consideration all the economic issues, as well as the availability of -- you know, all of the things that should be factored into their determination that that dose is ALARA, if you will.

MEMBER LEITCH: So it's more an assessment of does the licensee have a good program. Is he asking all the right questions?

MR. PEDERSEN: Correct. Now there's two aspects. When you compare the actual dose, collective dose that was experienced from a job to what was planned, if there's a discrepancy there, that could be from two different reasons. Either the planning process isn't very good, or the implementation of that plan isn't very good, so there's a performance aspect on both sides of that.

MEMBER LEITCH: Sure.

MR. PEDERSEN: If that happens, that's what we need to go in and look at. That's why we feel that additional inspection from the NRC or additional oversight is warranted.

Okay. Let me jump to -- the next two slides are actually not in your package. They're slides that I pulled out of an EDO briefing that are, I think, succinct summary of what happened at Callaway specifically.

CO-CHAIRMAN SIEBER: You'll have to provide us with copies of these.

MR. PEDERSEN: Yes, we will.

CO-CHAIRMAN SIEBER: Okay.

MR. PEDERSEN: I'm sorry I didn't. I didn't realize I was going to have to cover this in 15 minutes.

The fall of `99 outage at Callaway was very challenging to them. In shutting down, they had a CRUD burst that they didn't anticipate which caused the dose rates around the plant to go up significantly. They made a number of decisions as to what to do about that CRUD burst, and what to do about the jobs that were planned during that outage that resulted in significant discrepancies between what they considered ALARA in their planning process, and what they actually achieved.

CO-CHAIRMAN SIEBER: I have a short question.

MR. PEDERSEN: Yes.

CO-CHAIRMAN SIEBER: Callaway is a PWR?

MR. PEDERSEN: Yes.

CO-CHAIRMAN SIEBER: And typically you induce a CRUD burst when you shut down.

MR. PEDERSEN: Yes.

CO-CHAIRMAN SIEBER: Did they not do that?

MR. PEDERSEN: The details are fuzzy since it's been a few years. It's my understanding that they were trying a new process to induce the CRUD burst, and the CRUD burst didn't work. They got the CRUD burst at the wrong time. It wasn't being cleaned up as fast as they had anticipated. They decided to start the work without the CRUD burst being cleaned up.

CO-CHAIRMAN SIEBER: Okay. You can go on.

MR. PEDERSEN: As I said, there were a number of decisions that were made that were contrary to the ALARA planning that they put into place.

In fact, this is a list of the decisions of the issues that were brought out in the Notice of Violation that was issued for Callaway. They conducted work activities prior to the RCS cleanup and that affected a couple of jobs. They conducted activities prior to flushing the drains, et cetera. You can read them faster than I can talk about them.

They resulted in, on the next slide, actually three white findings. Earlier, I said two. I've misspoken. There were actually three white findings at Callaway in the ALARA area from this outage. The first white finding had to do with the scaffolding. As I said, we're judging their performance based on a unit of ALARA planning. At Callaway, as many licensees, their entire scaffolding, erection of the scaffolding is one planning unit. Now we call it one job. They pointed out that there were multiple, I think it was 57 JCNs associated with that, so they were trying to say that that was 57 jobs as opposed to one job, which we -- it was one of the points of contention in the appeal.

This action was appealed all the way up through the EDO, which is probably why we're talking about it as one of the issues as to whether we have the process calibrated properly.

Anyhow, the first job activity was the scaffolding. The first number there, the 22 person-rem was estimated. That was their planned ALARA dose. They achieved 46.35, a difference of 111 percent. The second job had to do with steam generator activities.

MEMBER LEITCH: Would you still have had the concern had they had this unexpected CRUD burst,

and went back and did some re-ALARA planning, and said hey, we've taken another look at this job. We've analyzed it. WE've done some things, and we now think the job is going to take 45 person-rem.

MR. PEDERSEN: That's exactly what we would expect them to do.

MEMBER LEITCH: Yeah. Right.

MR. PEDERSEN: And had they done that, none of these findings would be on this slide.

MEMBER LEITCH: Yeah.

CO-CHAIRMAN SIEBER: The other thing is to wait a little bit until --

MR. PEDERSEN: Well, that's the other thing too. Yeah, they could have just --

CO-CHAIRMAN SIEBER: -- the filters and demins absorbed the CRUD burst.

MR. PEDERSEN: There were a number of decisions that were made, and I won't go into my opinion as to why they were made, but --

CO-CHAIRMAN SIEBER: They all raised --

MR. PEDERSEN: -- they all impacted the dose, and there was no re-evaluation of what was ALARA for any of these jobs.

CO-CHAIRMAN SIEBER: Let me quickly ask another question. The steam generator work, that was

the electrosleeving work at that outage, or is this routine?

MR. PEDERSEN: I don't remember. I could --

CO-CHAIRMAN SIEBER: You don't remember.

MR. PEDERSEN: No. The third finding is something that I guess I didn't explain clearly. Could we go back to the actual SDP slide? The criteria for going to a white finding, there are two paths to that white box at the bottom. One is if an individual activity exceeds 25 person-rem, there's a performance deficiency that's made it through the screening process. In other words, the performance deficiency was greater than minor. It exceeded their planned ALARA dose by more than 50 percent. It was greater than five person-rem, et cetera.

If that resulted in greater than 25 person-rem, that's a white finding by itself, a single individual. And that's the first two of these. Collective dose, the nature of collective dose being the sum of many smaller doses, it was also recognized that you could have a significantly, or excuse me, a significant impact on collective dose, the overall performance of the program by having multiple failures of the program that don't exceed the 25. And that --

CO-CHAIRMAN SIEBER: Does this mean that the object behind that is to cause greater refinement of what a job is, so as to have smaller increments of dose? That's how you would defeat an absolute number.

MR. PEDERSEN: That is a safeguard built into it. That wasn't the rationale that went into providing for that path to a white finding. The rationale was that if you have, you know, five jobs that are greater than 5 person-rem, and you've had program deficiencies in all five of those, that that is exactly the same as having a program deficiency that has a 25 person-rem impact on your collective doses.

CO-CHAIRMAN SIEBER: Okay.

MR. PEDERSEN: That was the rationale, early rationale. What it does, however, is it does prevent you from saying well gee, if I plan all my jobs down to one person-rem, then I don't ever have to worry about getting through this process.

CO-CHAIRMAN SIEBER: Right.

MR. PEDERSEN: Which is an issue that came up when we discussed in public meetings. That's my presentation, I believe.

MEMBER ROSEN: I do have a question about the second and third. Now are those the same

activities, that you got them twice on?

MR. PEDERSEN: No. The first activity is scaffolding erection.

MEMBER ROSEN: No, no. The second and third.

MEMBER SHACK: Steam generator activities.

MR. PEDERSEN: Oh, the steam generator activities that are there in the --

MEMBER ROSEN: Yeah.

MR. PEDERSEN: No, they're separate activities.

MEMBER ROSEN: HP supports steam generator activities, they're not part of steam generator activities?

MR. PEDERSEN: They were planned separately. They were identified as separate units.

MEMBER ROSEN: No double jeopardy here. Now that's against the law.

MR. PEDERSEN: Right. We tried not to build that into this process.

MEMBER ROSEN: You can only be tried for a crime once.

CO-CHAIRMAN SIEBER: Yeah, but this isn't a criminal case.

MR. PEDERSEN: But to reiterate, we

started this process, the SDP process, with the understanding that it was to try to inform the assess process of the significance of an individual inspection finding. We didn't ever try to correlate between cornerstones. There was no attempt to determine how many person-rem collective dose corresponded to whatever conditional core damage frequency. It just wasn't in the process.

MEMBER LEITCH: May I ask you a question about that third sub-bullet, foreign object search and retrieval. That sounds like something that's evolved during the course of the outage, and I don't know whether it was or not, but it kind of sounds that way.

MR. PEDERSEN: No. I think they had a problem with it prior to that, and so they actually planned for that job.

MEMBER ROSEN: That's a fairly standard activity. Reactor Vessel Work, FOSR they call it, foreign object search and retrieval.

CO-CHAIRMAN SIEBER: Well, looking for it is standard, but trying to get one out that's lodged in there may not be standard.

MR. PEDERSEN: Yeah. That's the retrieval --

CO-CHAIRMAN SIEBER: You may be cutting

holes.

MEMBER LEITCH: Yeah, the retrieval part is what -- let's just assume in a hypothetical case that you had a foreign object that you're trying to get out, and you do some initial ALARA planning, and you say one and a half person-rem. And you use the one and a half person-rem, and you still don't have it out. You go back to do more ALARA planning, say you've got to spend another two person man-rem to get this thing out. Does that kind of an activity give you a problem?

MR. PEDERSEN: No. Actually, that's what we expect.

MEMBER LEITCH: That's what you expect.

MR. PEDERSEN: In that re-evaluation, however, we would expect the licensee to have a better idea as to what the cost in terms of man-rem was going to be, balance that against other consequences of maybe leaving it in there, or other remote handling. Whatever could be put into place to reduce the doses that weren't justified by the original cost.

MEMBER LEITCH: Yeah.

MR. PEDERSEN: The original was one man-rem. Well, we could just go in and grab it and pull it out. Now it's going to be ten man-rem. Well,

maybe we should have an engineer to remotely try.

MEMBER LEITCH: Right.

MR. PEDERSEN: Whatever those decisions are, and we're not trying to -- again, we're not trying to second-guess licensees.

MEMBER LEITCH: You're not willing to willy-nilly go from one and a half to six.

MR. PEDERSEN: That's right.

MEMBER LEITCH: You have to stop in the planning, reassessing the situation.

MR. PEDERSEN: Exactly.

MEMBER ROSEN: Now let me see if I understand what you're saying. If at the end of one and a half man-rem they still didn't have the object out, they knew where it was, and they knew what it was, and they had stopped the job and gone back and said here's what we're going to have to do to get it out. It's going to take us another five man-rem. We're going to have another job, because we're going to have to do a bunch of different things than we were doing. A five man-rem job, and then they went in and did it, and ended up with a total of 6.39 man-rem, then they wouldn't have had -- that wouldn't have appeared on the slide. Is that correct?

MR. PEDERSEN: That's correct.

CO-CHAIRMAN SIEBER: Well, you end up with two jobs.

MEMBER ROSEN: You end up with two jobs. The first one within the --

MR. PEDERSEN: Maybe. Sometimes some licensees would initiate a different ALARA package. Some licensees would just use the same ALARA package, call it the same job, and re-evaluate what the man-rem they expected, and come to that determination that it is ALARA to do that. It's the licensee's process that makes the ALARA determination. We're not second-guessing those decisions, unless they're obviously unjustified. But the requirement is for the licensee to have a program to implement engineering controls and procedures to minimize the doses, if necessary. That "if necessary" is a very subjective issue, and we've left that to the licensee's program to decide. If the licensee is running an adequate program, the outcomes of that program is what we're judging their performance against.

MEMBER LEITCH: And none of this involves individual over-exposures.

MR. PEDERSEN: No, there's a whole second half to our SDP in the occupational cornerstone that talks to individual over-exposures.

MEMBER LEITCH: You could go through all of this without over-exposing any individual.

MR. PEDERSEN: Right. That's correct. As a matter of fact, it's not up there. The flow chart, if you notice, only went to a white finding. The original flow chart didn't go passed yellow. There was an early recognition that ALARA issues would not take you to a red finding. The only way to get to a red finding in our cornerstone is an over-exposure for an individual five times the dose limit, significant over-exposure, and that's a red finding. That's the only way you get to red.

CO-CHAIRMAN SIEBER: Have you made any attempt to correlate the risk, mortality risk due to ALARA at your limits here, versus an early fatality risk related to CDF?

MR. PEDERSEN: No. I've made no attempt to do that. There's a number of difficulties built into that whole concept. First of all, collective dose, if you blindly take the linear no-threshold hypothesis as gospel, you can calculate numbers.

CO-CHAIRMAN SIEBER: Right.

MR. PEDERSEN: There's a lot of uncertainty when you extrapolate down below 10 rem, whether what you calculate means anything or not.

CO-CHAIRMAN SIEBER: Well, there's new opinions coming out all the time.

MR. PEDERSEN: There's a lot of -- yeah, you're right. There's a lot of controversy right now in the radiation protection business as to whether the linear no-threshold hypothesis should be extrapolated all the way down to virtually zero, which is what --

CO-CHAIRMAN SIEBER: It's the latest.

MR. PEDERSEN: It's not the latest by this agency, but there --

CO-CHAIRMAN SIEBER: It's the latest I read.

MR. PEDERSEN: There are other agencies that have published risk factors down to per Becquerel, per disintegration per second, per 100 square centimeters of exposure. We've had some difficulty with that. We, the NRC, provided some comments to that, but that's a whole other issue. So even if you're talking about individual --

CO-CHAIRMAN SIEBER: Well, given that you need --

MR. PEDERSEN: Well, even if you're talking about individual exposure and not collective dose, you're talking about dose to the work force, and how do you compare that with the potential dose to a

number of the public through CDF? Even if you go to a full level 3 PRA you're talking about how do you balance the dose to the public versus dose to the occupational worker. There's a lot of issues in there that are very difficult --

CO-CHAIRMAN SIEBER: Well, the source term -- okay. Well, let's move on from there. I'm sorry I asked.

MR. FRAHM: Thanks, Roger. Next we have public radiation safety. Steve, are you ready to talk about it? Steve is, I believe, under the weather today, so be easy on him. We have ten minutes before the break, and we actually have two specific examples we wanted to go through, so I guess optimistically I hope we could get one before the break, and maybe pick this up right after the break, and then move into emergency preparedness.

MR. KLEMENTOWICZ: Good afternoon. Yes, I'm a little bit under the weather recuperating from a cold over the weekend. The public cornerstone, the overview is that it's designed for routine plant operation where radioactive material is either released into the environment, transported into the environment, or inadvertently brought into the environment. It's made up of four branches,

radioactive material control, transportation, radioactive effluent control, and environmental monitoring programs.

The example we'll be talking about is radioactive material control. The issue was with Comanche Peak. In the SDP, we have a sub-routine. I could point it out on the overhead here. We have a sub-routine that talks about how many occurrences, and that's how many occurrences over a two-year inspection period.

The public cornerstone deviates from some of the other cornerstones because besides being performance based and trying to be risk informed, we also have a public confidence factor. Because this cornerstone involves the public and radioactive material in the public domain, as one of the agency's goals, performance goals and objectives, public confidence is something that we are to promote. And any time radioactive material gets into the public domain, we know that the public is greatly concerned about that. So even though we do have dose standards and dose limits, and ALARA objectives for effluents, the public confidence factor was put into this cornerstone with the agreement of industry and stakeholders, public stakeholders.

Industry itself recognized that they would lose tremendous public confidence if they were to release material inadvertently, as we saw in Davis-Besse, and so the industry agreed, and we felt it was appropriate at the NRC to have this public confidence factor. So that's a major difference that I need to point out to you, that we have this public confidence factor that is subjective. We try not to build it up to such great extremes where it becomes an outrage factor, as has happened many years ago, but it is there.

That's where -- partly what this greater than five occurrence loop was to consider. It was also to consider that if you had very low level material releases on workers or contaminated soil or equipment -- I had just come off the Haddam Neck assignment where they had released contaminated blocks in soil throughout the countryside, and what we found there was there was multiple very, very low level doses from each one of these concrete blocks.

MEMBER WALLIS: What's the threshold for these very low doses?

MR. KLEMENTOWICZ: Five millirem.

MEMBER WALLIS: Well, that seems to be a bit more than five millirem if you go the other way.

MR. KLEMENTOWICZ: Correct. We have two branches. One is strictly dose-based. The other we add occurrences.

MEMBER WALLIS: There seems to be no threshold for an occurrence per se. I mean, is microrem an occurrence or --

MR. KLEMENTOWICZ: Currently, we have no release limits in 10 CFR Part 20. Many years ago, the agency tried to do a below regulatory concern to establish a threshold, and that went down in flames. The policy, the NRR policy is no detectible licensed radioactive material can be released other than effluents. So what we have is that the licensee has to have a material survey and release program, and it's based on instrument sensitivity, and so that becomes the de facto release limit. However, as I said, if it's ever detected, then that is a potential violation.

MEMBER ROSEN: Now wait a minute. I came into the plant with potassium 40 in my body.

MR. KLEMENTOWICZ: That's why I make it --

MEMBER ROSEN: Can I take my own potassium 40 back out?

MR. KLEMENTOWICZ: That's why I make it clear it has to be licensed radioactive material. It

has to by byproduct material that came from the plant.

MEMBER ROSEN: Okay.

MR. KLEMENTOWICZ: Everything is measured above background, or licensed material. So since Part 20 has no release limits, and I'm also on a working group where the Commission has directed us to establish a clearance rule, we hope that in the next several years we may have a limit. But currently we don't. We have a no detectible policy and, therefore, if anything is released and found off-site and is detectible, it's a potential violation.

At the Haddam Neck event, we found that these multiple events did not contribute a 5 millirem exposure, so what we would have as a situation, was all of this material was released over multiple -- over different time periods, and the public confidence would go down. And yet, all we could say this is a green issue, so we came to the number five that if there were very small releases, but yet it occurred greater than five times over two years, we felt that was worth a white finding, escalated NRC attention. And the example we have was Comanche Peak. They had eleven instances where they inadvertently released licensed radioactive material, and so they tripped the greater than five, and it became white.

The licensee appealed, and this appeal went all the way up through the EDO. And I have this example up here because it shows where we worked with stakeholders to refine the program. And this brings up your question about isn't there some de minimus level, and we've been meeting with stakeholders pretty much every month for quite a while now to try to establish what is a minor inspection violation. Give the licensee some credit that when they do surveys, they can only see to a certain level, and realizing if you want to account for 24 hours, you could see anything.

MEMBER WALLIS: It's the way you measure it too, if you're discharging something into the river.

MR. KLEMENTOWICZ: Correct.

MEMBER WALLIS: You have to measure it before it gets too dilute.

MR. KLEMENTOWICZ: Right. But here we're talking about workers carrying tools or equipment outside of the restricted area.

Comanche Peak felt it was completely unfair that some of their items, contaminated glove liner stuffed underneath a cap in a welding tank, a contaminated wrench inside of a tool box, and most of

this material was found within the protected area. And so they argued the public cornerstone's objective discusses things being released into the public domain. And here was a situation where there was negligible risk to the members of the public from this material being on-site and discovered. So they argued that, you know, you're not meeting your objective.

The NRC agreed with that philosophy that if it's within the protected area, then we should not be aggregating these findings to a white finding. So as of November 29th, the SDP has been changed to reflect that if material is found within the protected area, it will not be aggregated to a white finding.

CO-CHAIRMAN APOSTOLAKIS: We were told earlier that the colors were determined by the action the NRC staff would take. You didn't mention any action. You just talk about public confidence. Are you the exception?

MR. KLEMENTOWICZ: No, we do the same thing. As a result of the white finding at Comanche Peak, there was a supplemental inspection that went back to

the --

CO-CHAIRMAN APOSTOLAKIS: But that's not how you determined white.

CO-CHAIRMAN SIEBER: No, that was a result.

CO-CHAIRMAN APOSTOLAKIS: That was a result.

MR. KLEMENTOWICZ: That was a result, yes.

CO-CHAIRMAN APOSTOLAKIS: We were told that white is determined by the action, and you don't seem to mention that at all. You just go with millirem.

MR. KLEMENTOWICZ: Well, that's part of the performance in risk-based. The 5 millirem is equated -- yeah. Well, I'm not sure I fully understand, but the action that the NRC will take, but we developed this criteria that would trigger a white finding, and then initiate the NRC action.

MEMBER WALLIS: Do you have an action matrix the way they do with the other --

MR. COE: Sure. Absolutely. It feeds the action matrix just like any other finding.

MR. PEDERSEN: Could I add something that might help? As I said, the way we developed the threshold was by subject matter expert and industry, with industry and stakeholder input as to what action would be warranted at certain levels. Those levels that would warrant NRC addition inspection, what

caused that additional inspection is what Steve is talking about. How much the staff and the stakeholders factored in public confidence, how much we factored in safety in the occupational radiation area, the fact that one over-exposure is one of the metrics in our strategic plan, and we would have to report to Congress. All of those things factored into what level of response we would expect the NRC to be in for any particular of these issues.

MR. COE: Another way of asking the question of Steve, I think, would be does the NRC feel comfortable that a white level of response and effort is matched appropriately to this threshold?

MR. KLEMENTOWICZ: And the answer based on our stakeholder meetings is yes, based on the possibility --

CO-CHAIRMAN APOSTOLAKIS: I really don't understand this process. Here you're telling us when we first did it, we considered the release of material anywhere. Then the licensee complains. It's okay. If it's within the protected area, it doesn't matter, so we don't include that.

MR. KLEMENTOWICZ: Well, it's a learning process, and when we first developed it, we were being, I guess, overly conservative. And based on

something like two years of experience, we felt that we were -- we found that we were -- this is what came out of some of the public meetings on the Comanche Peak and the appeal process, that we were unnecessarily causing public concern by escalating an issue that had zero risk to members of the public, so we were doing -- we gave it an unintended consequence by telling the public that this was a white issue, when in fact all of this material was in the licensee's protected area and had no risk to them. So if you want to call it this way, I screwed up by putting it in the first time. And then we did not want to alarm the public unnecessarily.

CO-CHAIRMAN APOSTOLAKIS: Did anyone ever ask you when you are in the white area, that's the same as if you had X number of scrams per year.

MR. KLEMENTOWICZ: No. No. I could not equate myself to reactor scrams.

CO-CHAIRMAN APOSTOLAKIS: You could not relate it.

MR. KLEMENTOWICZ: Right. We do not have a PRA like that. Absolutely not.

CO-CHAIRMAN SIEBER: By the way, even though you don't issue a color because there is radioactive material outside the radiologically

controlled area, but inside the protected area, that material still has to be identified, marked, and if it has loose surface contamination, has to be packaged.

MR. KLEMENTOWICZ: It can still be a finding.

CO-CHAIRMAN SIEBER: It's a finding. This doesn't have any color.

MR. KLEMENTOWICZ: No, it has -- it can have a green color, but if it's outside of the protected area, of it's in the public domain, it will get at least a green color. Plus, it will be added in this counter. The only thing we modified was for findings that are within the protected areas, we would not add them.

CO-CHAIRMAN SIEBER: And so all the rules on packaging, marking and all that other stuff still stands.

MR. KLEMENTOWICZ: Still stands, yeah.

CO-CHAIRMAN SIEBER: Okay.

MR. KLEMENTOWICZ: All the regulations, licensee following their procedures still stands. We can have a finding. It would be a green finding, but if they a hundred, it's one hundred green findings as opposed to going white. And we -- again, the mistake we made up front was that the unintended consequences

of alerting the public to something that was not a risk to them, so that's where we agreed with industry that we needed to change that, so that's the significance of the Comanche Peak item.

CO-CHAIRMAN SIEBER: Well, we'll need to really accelerate ourselves right now.

CO-CHAIRMAN APOSTOLAKIS: Is this an SDP or a performance indicator?

MR. KLEMENTOWICZ: It's an SDP. I have performance indicators on radioactive effluents based on how much gaseous and liquid effluents they discharge.

CO-CHAIRMAN SIEBER: Is there a way we can sum up, because we were supposed to --

MR. KLEMENTOWICZ: The next item, and I'll finish up very briefly, is radioactive material transport.

CO-CHAIRMAN SIEBER: Yeah, and let's not do that, because we have to take a break until 2:45. Okay. We'll come back at 2:50. Thank you.

(Off the record 2:07 - 3:04 p.m.)

CO-CHAIRMAN SIEBER: Well, we're a couple of minutes late. We'll get started anyway.

MR. KLEMENTOWICZ: Okay. Welcome back. I'd like to talk about one of the other branches of

our public cornerstone, and that's the transportation area. The example that I have here is about a failure to properly classify radioactive waste shipments.

This event came right out of the starting gate of the ROP. The program took effect in March or April, and then this was at Peach Bottom. And let me show you the actual SDP. It's the low level burial ground SDP. And we go through the was it an access denial situation? Yes or no? In this case, the event was the Part 61.55 waste under-classification gate. The licensee had packaged -- had labeled the material Class A waste, when in fact it was B, so under the old SDP, this is the revised one you see here, any time a licensee under-classified a waste shipment, it would be an automatic white finding. And that's what we had worked through with industry based on the regulations, public confidence, and any risk to members of the public or to workers.

The white finding was issued, and the licensee appealed. And the basis for their appeal was while they did call the material a Class A shipment, they had packaged it and did all the transportation, and shipping, and packaging requirements as if it was waste of Class B waste. So when you went through the SDP, okay, the under-classified it. They called it A,

but they met all the requirements of Class B, so the licensee made the argument that there really was no risk to members of the public, or to the workers during the transportation or the burial, because the waste was adequately packaged. The only error was that it was mislabeled.

We looked at that, ran through various scenarios, and ultimately agreed that the SDP needed to be a little more complicated, and not so simplistic as to just say under-classification, automatic white. So we added in this box, "Did the waste conform to the regulations, the de facto performance-based criteria?" They may have mislabeled it, but was the waste properly packaged and transported?

CO-CHAIRMAN SIEBER: Question, part of the packaging and shipping is radiation survey of the package. Was that correct?

MR. KLEMENTOWICZ: That's correct. And that's where the error was. They made -- the finding was that they -- their data showed that this material was Class B waste, but they didn't believe their own instruments, and some inexperienced technician had run the analysis. So while everything was telling them it was Class B, the data came out, the computer program came out calling it Class A, so we were -- it was a

somewhat difficult decision because we were concerned about the negative effects that they just happened to over-package it. We did not want to give credit for luck, so that's why we expanded the SDP, that some licensees we found out through investigations conservatively package. While they believe it's Class A waste, they will package it as Class B just to be sure. There's that added conservatism.

CO-CHAIRMAN SIEBER: It seemed to me, and I may be wrong on this, but the amount of money you pay to Barnwell, or Hanford, or wherever you're sending it depends on what the waste classification is. Right?

MR. KLEMENTOWICZ: Yes. Correct.

CO-CHAIRMAN SIEBER: So they got a discount misclassifying it as A, when it should have been B.

MR. KLEMENTOWICZ: Well, when it's received at Barnwell, they do their own independent surveys.

CO-CHAIRMAN SIEBER: Yeah, I know how it works.

MR. KLEMENTOWICZ: Yeah. But those are things we were concerned with. But the bottom line, we had to look at the regulations, and if the proper

packaging for Class B waste was met, and it was disposed of at Barnwell in a Class B trench, then there really was no risk.

CO-CHAIRMAN SIEBER: Well, how did Barnwell put it in a Class B trench?

MR. KLEMENTOWICZ: Well, they did the survey.

CO-CHAIRMAN SIEBER: Oh, this is after they surveyed it.

MR. KLEMENTOWICZ: They did the survey and they say no, this is a Class A waste.

CO-CHAIRMAN SIEBER: Okay.

MR. KLEMENTOWICZ: So this was found out when it got to Barnwell.

CO-CHAIRMAN SIEBER: Okay.

MEMBER KRESS: Once again, this is a question of whether or not we should ever have risk as part of the equation because, you know, just the fact that they misclassified it as a performance issue, and just because it wasn't very risky, transportation in general is not very risky. And, you know, it seems to me like it's a performance issue, and it shouldn't be ameliorated because of the risk-significance of it. It's a performance issue. You don't want waste to be misclassified it, whether they packaged it right or

not.

MR. KLEMENTOWICZ: Well, it still is a finding in the future. Under the original SDP it was a white finding. Now the same situation would still be a finding, it would be a green finding, so it --

MEMBER KRESS: I think it should still be a white finding is what I'm trying to say.

CO-CHAIRMAN SIEBER: You're too tough. It's like a parking violation. If you park in front of fire plug and there is no fire --

MR. KLEMENTOWICZ: Well, the way we have it is that it will be green in the future. The exception to that is when we get to the higher class waste, Class C. Then, you know, that's risky material, and we're not going to give much flexibility on that. That would be a white finding.

CO-CHAIRMAN SIEBER: But that stuff usually goes in the HIC.

MR. KLEMENTOWICZ: Correct.

CO-CHAIRMAN SIEBER: They're pretty easy to pick out.

MR. KLEMENTOWICZ: Yeah. But again, that was the part. That's the higher activity material. We were not going to de-escalate that in any way. But on the basis of our program being risk-informed, if

there was no risk, then we should not be taking additional action on the licensee when there was no risk, so that's why we agreed to do down to a green.

To make it the higher classification, the highest we have is yellow, and that's where the public confidence comes in with, if the licensee has several green or white findings, what will typically happen is the burial site becomes very agitated, and they say you've made your last mistake with us. You are now banned from disposal, you know. And here's where public confidence comes in.

The industry and the stakeholders agreed that while that may not represent the true risk to anybody, it's a severe public confidence issue that a licensee of the NRC got banned from waste disposal because of multiple errors. And that's typically what it takes, multiple repetitive deficiencies where the burial site says we don't want your thousands of dollars per cubic foot. So in that instance, it would be a yellow finding, and that's our highest level for the burial ground activities. It results in a suspension for greater than 30 days based on multiple findings, and that has not been changed, nor has that been tested.

MEMBER SHACK: When we just have findings,

do we trend the number of findings for a licensee? Is that --

MR. KLEMENTOWICZ: Trends? They're reported in the ROP database, but as far as a trend, the policy if it's a green, it's a green, it's a green.

MR. COE: We have about six to eight hundred findings per year total out of the entire program. About two dozen, about 25 or so, get looked at as potentially greater than green, and about half of those turn out to be greater than green. And that's a rough average based on experience to date.

MEMBER KRESS: Does the fact that you have no red color in this area give the message that you don't think transportation is as important as the other cornerstones?

MR. KLEMENTOWICZ: No. This transportation is broken up into several sub-branches to take care of all the different transportation regulations.

MEMBER KRESS: Oh, I see.

MR. KLEMENTOWICZ: We can get a red, and my next plan was to discuss where we could -- how we're dose-based, and we use dose for risk, to be risk-informed, and so we can get to red if they exceed

the regulatory limits. So clearly, if your multiples of the limits, you can go all the way up to red. We planned for situations that do occur. The package is breached. It's on the road, and we've had a few of those just recently where part of the material broke through the wall of the C-van, and now that's a package breach. But then we looked, are there any loss of contents, so we try to say what is the performance? Did any material leak into the public domain that could affect members of the public? Yes or no? If the answer is yes, then what were the dose consequences of this breach? So that's how we factored in performance with the regulatory limits and multiples of the limits. But again, if you exceed the public dose limit of 100 millirem, then that's going to get you a red. If you exceed the occupational dose of 25 rem, that would go red, so we've addressed public and occupational workers.

MEMBER KRESS: I think should avoid calling that criteria risk criteria.

MR. KLEMENTOWICZ: I should avoid calling it risk-informed?

MEMBER KRESS: Yeah, because there's no probability of frequency associated with it. It's all right to use it. I'm not against using it. Don't

mistake me. I just wouldn't call it risk --

MR. KLEMENTOWICZ: Right. And as far as reactor safety cornerstone, you know, we're totally different. But again, the concept is that we blended in dose, use of the regulations and public confidence to come up with a finding classification, and with this one exception of the under-classification, the rest of the cornerstone has worked very nicely.

MEMBER WALLIS: It's very interesting that you've mentioned public confidence many times.

MR. KLEMENTOWICZ: Yes.

MEMBER WALLIS: Who decides what the scale is for public confidence?

MR. KLEMENTOWICZ: That was based on our interactions with stakeholders.

MEMBER WALLIS: So you can actually poll the public or something?

MR. KLEMENTOWICZ: Based on all the meetings we had, we said what would be unacceptable? Where would the public -- it's subjective.

MEMBER WALLIS: Are these stakeholders members of industry, or are they members of the public?

MR. KLEMENTOWICZ: Both.

MEMBER KRESS: That's probably the best

way to get performance indicators when you looked at thresholds. I don't know of any other good way to get them. That's probably the best way.

MR. KLEMENTOWICZ: Well, let me go into this yellow example on the suspension. Now just because a licensee is banned from disposing of their waste from multiple minor infractions, that's a very subjective response by the burial site. You know, whatever -- when the governor gets too upset, he's going to say you're banned, so there's a subjective criteria right there. But what follows through? That's on the front page of the newspapers. The public reads Indian Point banned from waste disposal site for multiple violations. Was there any risk to people? Yeah, maybe slight, but it's a public relations nightmare. And what would be the expected NRC response? Clearly, green is not appropriate. They've just been banned. White, it did not seem significant enough. Yellow, we would have to find why are they doing repeat violations of this material that's in the public domain, and that is to be buried safely, and the public has to have the assurance that the waste is properly disposed of.

MEMBER WALLIS: What concerns me is you are making a decision of giving an award of yellow

based on somebody else's evaluation.

MR. KLEMENTOWICZ: Yes.

MEMBER WALLIS: But you're wrong.

MR. KLEMENTOWICZ: Yes, and that was a big discussion point, but the bottom line was that there were -- there would be violations. The NRC would have likely green findings, multiple green findings, so there would be performance deficiencies that would be documented. But the public confidence factor is what industry agreed that it was appropriate for the NRC to take additional action, because that would reflect entire industry.

MR. COE: I would offer that the SDP, as all SDP our staff -- they're defined ultimately, and approved and used by the staff. They benefit from the dialogue that Steve has talked about. And all of the SDPs have benefitted from similar dialogues with public and utility stakeholders. But when it's finally printed in an NRC inspection manual chapter and utilized by the staff in our decision processes, it is our decision process.

MR. KLEMENTOWICZ: That completes my presentation.

CO-CHAIRMAN SIEBER: Okay. Any further questions? If not, maybe should go home, get some

chicken soup and recover.

MR. KLEMENTOWICZ: I intend to. Thank you.

CO-CHAIRMAN SIEBER: Okay. All right. We'll just move right along.

MR. FRAHM: Thanks, Steve. Next we have Randy Sullivan to go over some emergency preparedness issues. Last but not least.

MR. SULLIVAN: I'm surprised. I didn't expect to get through this long agenda and be here. Hi, I'm Randy Sullivan. I'm a Senior Emergency Preparedness Specialist in NRR. I was the principal contributor to the EP cornerstone when it was being developed. I appreciate making a short presentation to you. I want to go to backup slide 31, and I want to begin there. It's a little different than maybe you were expecting.

This is a big surprise to us that there would be so many findings in EP. We've spent a lot of time wondering about this, studying it, trying to put it in context. We spent a lot of time --

MR. FRAHM: And actually, let me just point out, this slide looks a little bit different from the one in your package. And, in fact, there's an actual change. There's four white PI results in

the EP cornerstone versus three, so when I went and made that change, I also went and kind of columnized this slide to make it a little more legible.

CO-CHAIRMAN APOSTOLAKIS: Take the other one.

MR. SULLIVAN: Okay. Fine. Why don't I just take the other one. It started off easy, you know, because the early findings were Indian Point, and we could see that their program was a bit, perhaps had been -- not gotten the attention that you might have expected. But there kept on being findings, so we wanted to look at that, and we re-examined the SDP to see if we were in the right place. We examined the findings themselves. WE asked ourselves a lot of questions, can this possibly be equivalent across the cornerstones? You know, perhaps our view is myopic because we're EP experts, we're not reactor safety experts. And we came to several conclusions which I just want to relate to you.

Okay. So we have 20 findings in EP since the beginning of ROP. That's kind of a high number. There's five PI hits rather than the four you see on this slide. But there's some insight to be gained from this. The findings are grouped. We have three at Exelon, Pennsylvania; three at Indian Point; four

at Cooper, and three at NMC-Wisconsin. If you remove those numbers, you have seven findings over the other 57 sites. You know, that tells us maybe these programs were identified. You know, maybe the programs that racked up these findings were the ones that needed attention.

By the way, it's kind of instructive to note that of the five PI hits, I mean, PIs crossing a threshold, three of them are from that same group. Nice sort of -- when the -- the EP cornerstone is designed to identify problem programs, and to focus the effort there. If a program is operating in the green band, our inspection is more focused on problem resolution, critiques, rather than the performance itself. When a program ends up with these findings, then we get more involved with the performance.

Now the original EP SDP recognized that there could be false positives. We'll go through a little bit of the SDP, but that was actually written into the cover page of the EP SDP. It's a long paragraph that's kind of well-written, but it basically say we recognize that we could have false positives. This SDP was designed to have no false negatives, and what that means is that there may be at times be a finding that is characterized at too high

a level. In other words, it runs through the SDP as white. Maybe it should be green or yellow, maybe it should be white. And it gave the panel the latitude to use that judgment. It's supposed to be the exception rather than the rule. We think perhaps it was invoke more than - it was invoked I think three times, twice, three times. We think that was too many, so - but nevertheless, it was there.

We are in the process of almost finalizing a revision to the SDP that tightens up several areas. We think we learned over the first couple of years, and we did change the SDP to provide some flexibility. It was a little inflexible in terms of the risk-significant planning standards. I'll explain that concept in a minute or two. It was either yellow or green. We were kind of unsatisfied with that, so now there's an intermediary step of white, and it actually tightened up the critique finding to make sure it really is doing what we wanted it to do.

Okay. I'd like to move on to a couple of examples. When you take a look at emergency preparedness -- well, I'll tell you what. Rather than do that, why don't we look at the SDP? Can you put up the SDP? We thought this was simple. It looks kind of simple, but it's not.

You have a finding. There's three paths to go down. Actual event is the far left, we'll get that in a minute. If it's a drill or exercise critique problem, you go down the middle. If it's a risk-significant planning standard problem that wasn't ID'd, then it's white. If it's anything else, it's green.

We have 16 planning standards in emergency preparedness and some requirements in Appendix E. Rather happily, four of those planning standards relate most directly to protection of the public. That's classification, notification, PAR development, and assessment, dose projection and the like. So those are what we call the risk significant planning standards, because they live closest to protection of the public health and safety. Sirens are subsumed in notification, planning standard 5, so it's both notifying the off-site agencies and notifying the public are talked about in planning standard 5. That's what we got.

The other eleven planning standards are less important or less significant. You've got to comply with them but, you know, it might be a white finding instead of a yellow finding, or a green finding instead of a white finding. That's just the

way it is.

On the right hand side is the failure to meet. That's an actual programmatic deal. Failure to meet a planning standard, no, it's green. You know, some plan commitment or some other thing starts green. If it's a failure to meet a planning standard, you're at least going to get a white finding, and if it's one of these risk-significant planning standards, it's yellow.

On the next page is real events. Anything you do wrong in an unusual event can't be worse than green. There's about 30 unusual events a year. There's about three alerts a year, so some of the things you do wrong during an alert, like a failure to classify which is the Peach Bottom case you brought up this morning, can be white. Any of the missed steps that aren't associated with the risk-significant planning standards would be green. And it bumps up in that fashion for site in general. You can get to a red under the general emergency. Okay.

CO-CHAIRMAN APOSTOLAKIS: What's RSPS?

MR. SULLIVAN: Risk-significant Planning Standard, that's classification, notification, PAR development and assessment, 50.47(b)(v)N.

CO-CHAIRMAN APOSTOLAKIS: So if there is

a general emergency, and they fail to implement RSPS -

-

MR. SULLIVAN: That would be fail to classify or fail to notify, or fail to issue a protective action recommendation.

CO-CHAIRMAN APOSTOLAKIS: In a real emergency.

MR. SULLIVAN: Real emergency, not a drill.

CO-CHAIRMAN SIEBER: In a general emergency, that would be the least of your problems.

MR. SULLIVAN: It's the least of your problems, yeah. Nobody argued with this much because the next general emergency --

MEMBER ROSEN: The next guys are going to have to pay a --

MR. SULLIVAN: Yeah. It will be the last general emergency. Industry didn't really argue with that much, but that's our only red finding. And we think that's appropriate. We can get yellow findings, and we have gotten yellow findings, but a red finding in EP is really only if you really deny the locals a chance to protect the public. And that would have to happen, you know, during a general emergency.

Let's go to the examples. This is an interesting one. If you do a bottom-up analysis of emergency preparedness, you find much to our surprise when we did it, that the siren system is absolutely the most important piece of gear you've got in emergency preparedness.

Now under the old program we would invest quite a bit of inspector time looking at field monitoring kits, and equipment lockers at TSCs and the emergency lighting within the -- we don't do any of that any more. But we do look at the siren system because you cannot protect the public health and safety without the siren system, in fact. There's no other -- you know, it would be nice if you notify, but if the locals can't then light off the siren system and turn on the EDS station, you don't protect public health and safety, so we invented this PI that's a little unsatisfying.

We've been calling it reliability. In any case, it's a measure of successful tests over tests on a per siren basis. We took 60 plant years of data. We looked at the average. It was high. The regulatory limit is 90 percent. The average was 98 percent. We chose 94 percent for the limit. In the 60 plant years of data, there was one plant that was

below that, I mean, one data year that was below that, one data year that was near it. We chose it on a consensus basis, declared victory and pulled out. Lo and behold, we get these findings.

CO-CHAIRMAN APOSTOLAKIS: Why didn't you do what the safety guys did with their indicators, where they considered the plant-to-plant variability curve, and they took the 95th percentile?

MR. SULLIVAN: Good question. Maybe we weren't that sharp, but what we did do is we took an average which turned out to be 98 percent. Oh, I'm sorry. I shouldn't have answered you that way. This, in fact, is something like one sigma off the -- I mean, if you use sigma in a very loose --

CO-CHAIRMAN APOSTOLAKIS: Yeah, well they didn't do it that way. They actually went to the 95th percentile.

MR. SULLIVAN: No.

CO-CHAIRMAN APOSTOLAKIS: That's more than one sigma.

MR. SULLIVAN: Right. It would be two sigma. Right? We used one sigma for --

CO-CHAIRMAN APOSTOLAKIS: I mean, why not do what they did? It's not a matter of being smart. It's a matter of having somebody overseeing the whole

effort and saying, you know, this is how we do it. I think, you know, we keep talking about public confidence. I think we are undermining public confidence by doing things like that. In the same program, some things are done in one way, some other things in a different way, some other things in yet a different way. That's what we're -- you know, that's a major determinant of public confidence, in my opinion.

CO-CHAIRMAN SIEBER: That pretty much goes back to what we said, you know, months ago about consistency.

CO-CHAIRMAN APOSTOLAKIS: Yeah.

MR. SULLIVAN: We were constrained by the 90 percent reliability number. That's the FEMA regulatory number. If you drop below 90 percent, FEMA gets involved in your siren system reliability, so we felt that that was an absolute floor.

CO-CHAIRMAN APOSTOLAKIS: Yeah, but presumably then all 102 units are above 90 percent.

MR. SULLIVAN: They are above 94 percent.

CO-CHAIRMAN APOSTOLAKIS: Yeah.

MR. SULLIVAN: The average is 98.

CO-CHAIRMAN APOSTOLAKIS: Yeah.

MR. SULLIVAN: So they're well above it.

CO-CHAIRMAN APOSTOLAKIS: They're well above it.

MR. SULLIVAN: So rather than use the analysis you're talking about, we felt constrained by the 90 percent, and we did a much simpler analysis.

MEMBER WALLIS: How many tests do you run to get this 98 percent?

MR. SULLIVAN: It varies from site to site. They'll file a siren design document with FEMA which was formerly approved. In that design document is the testing regimen. In general, it's a bi-weekly test, so 26 a year.

MEMBER WALLIS: This is averaged for three years or something?

MR. SULLIVAN: No, it's averaged over on year, but it's on a per siren basis, so if you had 100 sirens, there's 100 siren tests every two weeks. And so the number get happy. Now many sites do a lot more than that. There are sites that test daily, so they turn in 4,000 tests per month. And actually, the PI is designed to encourage testing, because the more tests, the more stable the number is, and one missed step doesn't make you cross the threshold to --

MEMBER WALLIS: Do these sirens work in ice storms and things like that?

MR. SULLIVAN: Yeah. It is challenge though. Salt water ice storms are particularly challenging. In any case, we ended up with quite a few findings in this area.

In this particular item, this system was not one of the better ones. They started to trend downward and then they had a system failure, so they were probably only testing every two weeks. They were already at 96 or 95 percent. They had a total system failure, and it drove them down below the threshold. We got involved. They did a root cause analysis, and that's where we were. This is a PI, this is crossing a PI threshold.

The next item is a finding, and this is an interesting one. This finding has to do with -- let's see. Siren systems have gotten more sophisticated over the years, and the systems now have control units with feedback, so there's a radio at the siren that talks back to the central, and it says, you know, it gives health and safety data, or health and welfare data. Maybe seven data points, maybe 20 data points, whatever it is, so when my signal goes out for a test, the siren comes back and says I'm okay. Or it says I sounded, because sometimes there'll be a little speaker, you know, it's fairly sophisticated. So

there's a siren feedback system.

However, many of the older systems don't have that, and among those systems, there are some that have no way of knowing whether the siren sounded or not, so they have automatic route alerting. It's adequate. You know, the sirens are designed to work, they'll probably work. You know, you're giving me 98 percent reliability, so automatic route alerting may be necessary. They'll do it anyway. There's some benefits to that.

At this site, they had a feedback system that wasn't working and they didn't know it, and there was no automatic route alerting. They could do route alerting, but they didn't know to ask --

MEMBER ROSEN: What does that mean, "route alerting"?

MR. SULLIVAN: It's firemen and policemen run a route with a bullhorn and tell people to get out of their houses.

MEMBER ROSEN: That doesn't sound so automatic.

CO-CHAIRMAN SIEBER: Automatic is when you notify, and they go without being told to.

MR. SULLIVAN: Let me say it a different way. These fire trucks and policemen leave

automatically. They don't have to be told to leave. If there's a general emergency, they're summoned and they start their routes. That's what I mean by automatic.

At other sites with the feedback system, they say well, siren six failed. Get, you know, police car A to go run its route. That's all we need. At these less sophisticated sites with no feedback, they all go. As soon as they get to their police cars they go and they run their route.

By the way, route alerting is the way public evacuation works everywhere else in America. That is the way neighborhoods are evacuated should there be a tanker truck turned over, or a train derailed, is route alerting by police and firemen, so although it's foreign to use in the nuclear industry, that is the way the whole country operates in -- you know, where there's no siren system.

Well, so initially it looks like they're not implementing a risk-significant -- they're not meeting a risk-significant planning standard. They cannot assure about 100 percent of the people will get notified should the sirens be needed. They thought they had a feedback system. The feedback system in fact was not working. They were unaware of that.

They would push the button and have no idea. Well, you would assume most of them would work, but you could not assure that 100 percent, about 100 percent of the people were notified within 15 minutes, so we felt well, you're not meeting the planning standard.

Yellow just didn't seem right. And our SDP was inflexible enough that it was either yellow or it was green. Well, we used that judgment clause and declared it to be white.

CO-CHAIRMAN APOSTOLAKIS: How could the process be inflexible to go from green to yellow without going through white?

MR. SULLIVAN: Well, let me help with that. It's kind of simple minded. Can you put this back up, the first slide of the SDP? It looked like a good idea when we started, but in fact maybe it wasn't.

CO-CHAIRMAN APOSTOLAKIS: There has to be some continuity in the judgment.

MR. SULLIVAN: If you look at the right-hand side, the way this is rigged, it's a failure to meet a regulatory requirement.

CO-CHAIRMAN APOSTOLAKIS: Right.

MR. SULLIVAN: You drop down. Is it a failure to meet a planning standard? If the answer is

no, it's just some plan commitment, but not a failure to meet a planning standard. It's green. If you fail to meet a planning standard, you drop down. Is it a risk-significant planning standard? The risk-significant planning standards include notification of the public. That's 50.47(b)(v) .

CO-CHAIRMAN APOSTOLAKIS: Right.

MR. SULLIVAN: If you fail to meet it, it's yellow, period, no step for white. So when -- we've rewritten the SDP to put in an intermediate step. We'll call it a degraded risk-significant planning standard and it will give the SERP more room to assign a white when we think a white is correct.

CO-CHAIRMAN APOSTOLAKIS: So there was a problem with the original --

MR. SULLIVAN: Yeah, sure. Frankly, we thought these kinds of failures would be so rare, these systems were 20 years old. They had been out there. They had been reporting good data to FEMA, and when we looked at it closer, many of these findings have been in ANS, and they've been difficult to struggle with. We've taken those lessons and we've rewritten the SDP to help a bit more. I'm sure we'll still be challenged. That's what we're trying to do. I'm done, if you're done.

MEMBER LEITCH: There are a number of places that are spending big bucks replacing siren systems, and I wonder to what extent that's being driven by this process. In other words, if we say this is performance-based, not particularly assessing the risk of the situation but assessing performance, and I think particularly when you get in the area of emergency planning, the public interprets it as risk-based. And I just wonder if that's an unintended consequence, if the utility is really spending -- if we're forcing, forcing may not be exactly the right word, but if you're influencing the utility to spend really big bucks in an area that may not be -- where we may not be getting our bang for a buck. Not to say it's not important, but is it the most important thing we should be doing?

MR. SULLIVAN: Let me put this premise forth. It is the most important piece of equipment in EP.

MEMBER LEITCH: Yeah.

MR. SULLIVAN: So rather than buy me new field monitoring vans, or updating the TSC, or putting in a new phone line, I would rather see -- I mean, this is a revelation that ROP showed us. Yeah, we knew sirens were important, but (a) we didn't know how

many problems we'd find. And (b), we didn't realize they were the most important piece of gear until we actually sat down and did the analysis, so yeah. I think that's a fact of performance indicators. If you measure it, people will pay attention. And we decided this is worth measuring, and people are paying attention. And there have been problems revealed.

MEMBER ROSEN: I think that the reason that you're seeing that is that many localities rely on these sirens for evacuation, a natural phenomenon. And because of that, the towns and localities that the plants are situated in feel very strongly about the importance of these, not because of the nuclear emergency so much, although they recognize they'll be important in the nuclear emergency. They are grateful and pleased to cooperate and the rest if the sirens are upgraded and work better because of the alerting capability of the instruments for a natural phenomenon such as hurricanes.

MR. SULLIVAN: This stuff is 20 years old too, I mean, much of it. So it is, and many of these designs are no longer -- you can't get spare parts any more so it's not surprising that some of them are being updated.

MEMBER LEITCH: How do you feel about

operating -- I've been looking at operating event history on a daily basis. I'd say they were probably within the past two months, there's probably been eight plants that have their sirens totally crippled because of weather conditions primarily. Sometimes for a period of several days. I mean, major ice storms and so forth, so many of these siren failures where they're spending big bucks to correct, replace the siren systems can be fixed in half an hour, so what should be our reaction when the siren system is inoperable for 48 hours, total inoperability for 48 hours? I mean regardless of the cost, this is risk significant, would we not be very concerned when the siren is not working for 48 hours?

MR. SULLIVAN: Yeah. It's a dilemma. One way to approach it would be to change this PI to availability, and we're pursuing that. But it's successful tests over tests loosely called reliability. I'm told that's not the exact definition.

CO-CHAIRMAN APOSTOLAKIS: It doesn't mater. We will use whatever you like.

MR. SULLIVAN: Okay. Well, that's what we've been calling it. Well, in front of this scholarly body, I didn't want to be caught misusing

the word. But availability, we'd be happy with availability, so if you're above 94 percent, you fix the problems. Even if it's out for 48 hours, you use route alerting. I understand that in certain ice storms, route alerting could be challenged too, you know, but we're happy with the average availability of these systems, and they will be out for a day at a time.

CO-CHAIRMAN APOSTOLAKIS: So what you're calculating now is the failure of the sirens to start.

MR. SULLIVAN: Yes.

CO-CHAIRMAN APOSTOLAKIS: That's the PI.

MR. SULLIVAN: Yes.

MEMBER ROSEN: Start and run.

MR. SULLIVAN: Yes.

CO-CHAIRMAN APOSTOLAKIS: No, not --

MR. SULLIVAN: Well, it could be a silent test, which is less than satisfying too.

CO-CHAIRMAN APOSTOLAKIS: That's for the PI.

MR. SULLIVAN: That is for the PI.

CO-CHAIRMAN APOSTOLAKIS: For the significance determination process now, do you include the possibility of repairing it in half an hour?

MR. SULLIVAN: In the significance determination process, we've invented an algorithm that sort of bridges the gap between availability and reliability, and that's just being published now, so we've attempted to grapple with that. I don't know if we'd be down to a half hour, but we've attempted to put together an algorithm that addresses availability. And should a siren system be unavailable, yet the PI testing in the green, we might issue a finding, and it would be a finding against the program, you know, against maintenance.

CO-CHAIRMAN APOSTOLAKIS: The reason why I'm asking is in the safety, reactor safety SDP, I think recovery is considered routinely. Right, Doug?

MR. COE: Yes, where it's appropriate, recovery of the equipment is. And in order to meet the equipment's objective, yes.

CO-CHAIRMAN APOSTOLAKIS: Yeah. So we could do the same thing here. Now I don't know exactly how your algorithm comes with that.

MR. COE: Well, I'm questioning whether you can recover a siren in 15 minutes. If the objective is to notify within 15 minutes, it's going to be, depending on the situation, of course, pretty tough to recover that siren.

MEMBER LEITCH: Like for example, I'm familiar with a case where a utility went to actuate the sirens from the county and none of them actuated, and it basically was -- maybe you pressed the wrong icon on your computer. It's you don't click there, you click over here. You click over here and they all work, so I think that turned out to be a white finding, but it was --

MR. SULLIVAN: I'm sorry, sir. It's closer to what the chairman is saying. The icon was missing. It had been accidentally deleted from the screen and they didn't know it.

MEMBER LEITCH: Yeah, right. That's correct.

MR. SULLIVAN: Until the test happened, and that -- these sirens -- well, we've been through the mill on sirens, and it was a very sleepy issue three years ago. We had willfulness. You know, we had willful tampering of -- I mean, at two sites. Who would have thought that such a thing would happen. We've had these computer issues where an icon is deleted and, you know, who would have thought that that would happen. This just turns out to be a more important system and there are problems.

CO-CHAIRMAN APOSTOLAKIS: You said that

FEMA has 90 percent.

MR. SULLIVAN: Yeah.

CO-CHAIRMAN APOSTOLAKIS: So this occurs through other natural phenomena that require evacuation.

MR. SULLIVAN: FEMA has a -- well, you know that FEMA is responsible for oversight of nuclear plant off-site programs.

CO-CHAIRMAN SIEBER: Right.

MR. SULLIVAN: And part of those programs, the siren design-basis criteria are issued by FEMA, you know, so we use FEMA's determination in this. And if a siren system has a reliability of less than 90 percent -- now they use a calendar year. WE're using four quarters, it's regulatory involvement. FEMA will get involved.

CO-CHAIRMAN APOSTOLAKIS: It's like the EPA and the NRC in another context.

MR. SULLIVAN: I think we're closer.

CO-CHAIRMAN APOSTOLAKIS: But I thought that these sirens are used also in other emergencies.

MR. SULLIVAN: Yes, of course.

CO-CHAIRMAN SIEBER: Yeah, but other industries, like the chemical industry, is not required to have sirens.

MR. SULLIVAN: That's right.

CO-CHAIRMAN SIEBER: So the only industry that's required to have them is nuclear plant licensee, and it's under FEMA Reg. 1. Right? That's the regulation.

MR. SULLIVAN: Right.

CO-CHAIRMAN APOSTOLAKIS: So what other phenomena, I mean, if there is an earthquake or what?

MEMBER ROSEN: Hurricane.

CO-CHAIRMAN SIEBER: Or the chemical plant next door goes up.

CO-CHAIRMAN APOSTOLAKIS: No, but then you say they're --

CO-CHAIRMAN SIEBER: They're going to use the nuclear ones.

CO-CHAIRMAN APOSTOLAKIS: And they have plans for doing that?

MR. SULLIVAN: Sure.

CO-CHAIRMAN SIEBER: They do. That happened down in Louisiana some place.

CO-CHAIRMAN APOSTOLAKIS: So they're going to use it, but they have no responsibility for their functionality.

MR. SULLIVAN: The utility maintains them. The county operates it, and the county may operate it

for other purposes if it helps them.

CO-CHAIRMAN APOSTOLAKIS: Put it a different way. If there is no nuclear plant around --

MR. SULLIVAN: Then there's no sirens.

CO-CHAIRMAN APOSTOLAKIS: So if there is a chemical emergency there are no sirens.

MR. SULLIVAN: And FEMA has standards for sirens. I'm not familiar with them. Their nuclear siren standards are the ones that I'm relating to you. There are other sirens. You know, the county has a tornado siren in certain counties and, you know, certain hurricane alert along coastal areas. I'm not saying there's no other siren systems, but by and large when you see a public evacuation in America, there's no siren system covering it.

MEMBER WALLIS: I was listening. You seem concerned about whether or not the siren works.

MR. SULLIVAN: Yes.

MEMBER WALLIS: What assurance do you have that people hear it? I mean, audibility depends on lots of things.

MR. SULLIVAN: We actually learned some lessons there too. When the siren system is installed, there's a sound mapping verification that's part of the design basis.

MEMBER WALLIS: Weather makes a tremendous difference.

MR. SULLIVAN: Yes, it does. And what this has done, is it's done -- the time of year is then calibrated perhaps using an algorithm for winter and summer. There could be a gale blowing, in which case the siren might not reach its design sound. That's true, but we didn't design for the gale. We designed for 60 dB at the front door in normal, you know, winter and summer conditions.

MEMBER WALLIS: So and old person with a hearing aide not functioning won't hear it and things like that.

MR. SULLIVAN: Right. Actually, there's a study -- there's a whole set of case law that addresses that. It really is 60 -- FEMA would prefer that we only discuss 60 dB at the front door. We went into our administrative law judge --

MEMBER WALLIS: We can't even hear the grandchildren at 60 dB at the front door.

MR. SULLIVAN: Yeah. Maybe, 60 dB or 10 dB above background at the front door. In fact, there are sociological facts that cause for informal alerting networks. They really do exist. Is it a sociological fact.

MEMBER LEITCH: You may have a 98 percent success rate with getting the siren to work. The audibility may be down to 80 something.

MR. SULLIVAN: And, in fact, the 90 year old neighbor of your's, you will knock on her door and let her know that --

MEMBER WALLIS: She may knock on our's too.

MR. SULLIVAN: Yes, that's right. She may very well, and the neighbor you hate will be in the back seat of your car. I mean, these are just sociological facts. You know, they're kind of amusing, but in fact, there is informal route alert -- informal alerting, and neighbors just don't let neighbors stay behind. It just -- hurricanes, other events, that's just the way it works.

MS. WESTON: What about the hearing impaired?

MR. SULLIVAN: Yeah. The counties spend a lot of time on special needs groups, and so I -- you know, sometimes it's a shoebox with cards in it, but in general, it's a computerized system that's updated, you know, in accordance with their FEMA commitments of lists of people who have special needs. They'll be ambulances assigned and other workers assigned to pick

up on those special needs.

In general, in an evacuation time estimate, the special needs people take about the same time as the rest of the population. Rule of thumb, not always true, so the hospitals, and the jails and the deaf people really don't take any longer than the massive population to leave an area in general.

MEMBER WALLIS: The most dangerous is probably a discotheque, you can't hear anything.

MR. SULLIVAN: Yeah, but those are the young people and they're resilient anyway, so --

CO-CHAIRMAN SIEBER: If you're there, you're young, and if you're young, you're immortal.

MR. SULLIVAN: Well, thank you.

CO-CHAIRMAN SIEBER: Okay.

MR. COE: Mr. Chairman, if I can offer a summary comment. I think what you've heard today, and I hope we've achieved our objective of giving you a sense of why the staff in general feels comfortable proceeding as we have with the set of SDPs and PIs that are available and are in use.

I think what you have seen here is that in each cornerstone, subject matter experts have taken a fairly hard look with collaboration of industry and public stakeholders, and as well as our internal folks

to come up with a way of grading our inspection findings that could arise from our inspection activities in each of these cornerstones.

To the extent we can, we've used risk insights either on a general kind of industry basis, or on a plant specific basis where those tools are available. And in an ongoing process, or an ongoing manner, we continue to seek the inputs and experienced that we've gained as we have and as we continue to get, to make refinements, to adjust these thresholds to produce what we believe is an appropriate regulatory response for a specific finding.

We continue to see the aggregation of these findings on a unit-by-unit basis in the action matrix, and it appears to be providing a relatively good spread between the plants that get the most attention, and the plants that get exceedingly lesser levels of attention from us above and beyond the baseline program.

On this basis, we believe that the program is working, I guess as the slide here starts out saying. And I would be interested in the Committee's reaction to two things. One is, do you believe that we've accomplished our objective today. I'd be very interested to know of helping illustrate why we feel

comfortable where we're at. Secondly, I would be, of course, very interested if there are any remaining issues on your minds that we could better understand as a result of this meeting today.

CO-CHAIRMAN SIEBER: Well, let me say a few words about where we are, and where we've been, and where we're going. We wrote a letter back on October 12th, 2001 which you read, which is a lengthy letter that talked about a number of things, but among them were the inconsistencies that result in the differences between using PIs with colors versus SDPs with colors, and then equating those as though they were the same thing. And also, how we deal with multiple sets of colors. You know, two whites equal a yellow, two greens equal a white, that kind of stuff, and what the rules of the game were.

I think there were some telling things that occurred today. Of course, we elaborated on all that at great length, which caused the Commission to write an SMR that basically told the staff to resolve all these things, and consult with the ACRS in the process of doing so. I don't feel from reading that SRM that it's our obligation to necessarily write a letter. On the other hand, we just can't let this thing float off into oblivion either. Okay. And you

folks have invested a lot of time, and a lot of your own infrastructure and credibility into developing this program, so I'm sure that you want to try to keep going in the direction you're going without having to stop and redo a bunch of things.

So the questions becomes, are there irreconcilable differences? Are there things that can be done to remove inconsistences in some of these intellectual pitfalls that we seem to find ourselves jumping into from time to time to make the process seem intellectually more legitimate. And I think there are some telling things.

One of them was an observation by Steve Rosen where he defined what it is we think, what you think this process really is. And what it amounts to, and because of a failure of that fundamental definition I think, and the fact that we all don't see that definition as correct, is one of the root causes of the difficulties and the struggles that we're having, so I think that was one of the key statements that was made today, and should be taken into account.

I think that we would be remiss to allow this to float off into oblivion, and so we must think about responding to the Commission one way or another, even if it's an interim letter. And I think that you

folks have put a lot of effort into this process, and I think that it deserves at least a statement that there's progress being made.

I think that it is a living process and you'll never be done. No matter how -- as long as there's reactors out there and people making mistakes, I think there's opportunities to improve their corrective action system and our own. So I think that we're faced with the potential, since the Federal Register notice for the February meeting is already out, potential for us to ask you to come back in March so the full committee can further deliberate on what it is we want to do.

I think that in fairness now though, since we have a few minutes before we must close this portion of the meeting, that I ask our co-chairman here for his perceptions of what he's heard today, and how he puts this all together, and where he thinks we ought to go. George.

CO-CHAIRMAN APOSTOLAKIS: Sure. Well, first of all, before I go to that, I think when we have a process in place or a PRA in place, how do we decide that it's effective, realistic, or it's meaningful? Well, it seems to me the only way is to look at the real world, our experience with the real

world. And a process that gives green to Davis-Besse, and then we find out what's going on, just can't be effective. I don't know why you call it -- you say it's working effectively.

I mean, the fact that you are finding things at various plants, and then you miss such a major incident for a potential accident, in my view shakes up -- should shake up our confidence in the process. And we should really try very hard to fix it as soon as we can, because I don't care if I have 50 small things, siren here, or a transient here and I miss the big one, so I guess I disagree with you that the process is working effectively.

Now more generally, I'm not sure -- I get the impression that you really didn't take the ACRS letter seriously. Today we hear well, you know, we are considering abolishing the red. Then I pushed a little bit to say okay, we're not going to abolish it. And 30 seconds later somebody else says no, we're still considering it. Now that's not a serious position, you know, we are doing this, not doing that. I mean, it was very clear that you said that this is a meaningless number, and it has a fundamental flaw that you are determining it using the delta CDF by changing a single element of the PRA. And I didn't

see any -- you didn't address that issue, and we're getting conflicting -- I mean, at the end of January, after almost a whole year from the SRM, we're getting conflicting answers.

That tells me that in preparing to come here, you didn't really have a meeting and say how do we address this? This is the position, everybody say the same thing. And this is the reason for it. And I think you're more or less rejecting everything the letter said.

I haven't seen a single change in what you're doing as a result of that letter, so you're disagreeing with us. So maybe if we write a letter, we can write one line. We continue to believe what we said a year ago. Then we have to press to understand what the basic philosophical approach is. Is it performance focused? Is it risk? Then we get the answer that, you know, it's really performance. And I think from what you've described in certain -- the analysis of certain events, it is really performance with heavy doses of risk insights, which I think is great.

Then we ask, you know, how do you decide that these things are equivalent? And the answer was, based on our intended action, which now runs counter

to risk-informing the regulations, but also doesn't seem to be universally true, because later on when other colleagues of your's took the floor and they talked about other things, and I asked them did you decide these colors on the basis of action? They said no, so evidently there was not again a policy for the ROP that said look guys, this is how you determine yellow and white, based on what you would do.

I think that they were developed independently by various groups. We have another example with the sirens. They took the mean value and they went up a little bit. And here you have the performance indicators for reactors taking the curves from plant to plant, and using the 95th percentile. I mean why? Why can't they be consistent?

Is it going to make a big difference in what you're doing? No, but we keep talking about public confidence. The public is not just the average guy on the street. The public is also the statistical associations, the informed scientists. And if they take a look and they say well gee, these guys really don't know what they're doing, you know, that's not good. You're losing the confidence of important constituencies.

I believe that we should separate

performance from risk. And I think, you know, originally I was going to propose that maybe the reactor safety PIs should weight it more than say the emergency preparedness, because the emergency preparedness will be required after many very unlikely events occur. Whereas, if I have an initiating event, that really creates a lot of commotion immediately. But if I look at it from the performance point of view which Doug explained in the SDP for reactors, then maybe they should not be weighted, because as far as performance is concerned, if you don't do a good job in the emergency planning, it should be the same as if you don't do it in the mitigating systems.

So you see, if you have a philosophical approach, a lot of these things are resolved. If you say I'm performance focused with heavy doses of risk information.

CO-CHAIRMAN SIEBER: From time to time.

CO-CHAIRMAN APOSTOLAKIS: Where appropriate.

CO-CHAIRMAN SIEBER: Right.

CO-CHAIRMAN APOSTOLAKIS: Comma, where appropriate, period, as amended. So then, you know, the issue of the consistencies of colors, not so much whether white means the same everywhere, but should

two whites and a yellow be equivalent? Then this -- well, but this is not a problem with the ROP, but I think we need better guidance on root cause analysis, since a very important part of doing the SDP for findings is the determination, whether they're independent or not. And there is an underlying root cause, then it seems to me you have to be a little more formal when it comes to root cause analysis, and guide people, because you will only put their causes that come from your experience or your knowledge. And if you're not very familiar say with organizational factors, you never put anything there, unless it's obvious.

Then this other thing that Doug mentioned about timing, I'm a little uncomfortable with that. I can see your point, and again from the performance point of view, maybe what you're doing makes perfect sense. I'll have to think a little bit about it more, but from the risk point of view it doesn't. If it happened during preventive maintenance, well tough. The risk assessment will tell you this is a delta CDF. Right? But you are not risk-based. So you see, again, if you have a consistent philosophical point of view, it seems to me you will be able to resolve a lot of these issues and say, you know, this is the -- and

the other thing is the pressure boundary, it seems to me, is kind of unique, and you sort of agreed here, in the sense that, you know, one may still make the argument that if the cross-cutting issues are deteriorating, we will have advance warning. Say maybe a valve will fail here, or we'll see a consistent pattern of failures of hardware. But with pressure boundary, you may not have that luxury. And I think we need to pay special attention.

You may not have this advance warning. I mean, you had the steam generator rupture, tube rupture at Indian Point. And as we said, it was due to a defective inspection program, and then you have Davis-Besse, again defective corrosion control program. And we almost came close to an accident again because of the pressure boundary, although there there were indications. So again, this doesn't go back to the ACRS in all fairness. We didn't say anything at that time, but I think this is a new development now, and we probably have to pay more attention to this particular cornerstone as opposed to the other ones.

So that's what I -- oh, and the insistence of keeping the red in the performance indicators for reactors, and then two hours later we find that other

guys say well, we don't need the red. I mean --

CO-CHAIRMAN SIEBER: It doesn't make sense there.

CO-CHAIRMAN APOSTOLAKIS: Yeah. I mean, if it is logical not to have a red, don't have it. Why don't the other groups, I don't know which one it was now, why don't have this issue with public confidence? They certainly want to increase public confidence. Only the reactor safety guys feel that they have 24, 25 scrams there as a threshold for the yellow/red.

MEMBER WALLIS: Under some indicators you couldn't get to the red.

CO-CHAIRMAN APOSTOLAKIS: That's why we said that they should be abolished.

MEMBER WALLIS: It doesn't mean to say you abolish all reds just because for some indicators you can't --

CO-CHAIRMAN APOSTOLAKIS: No, for the transients you will never get there. You will never let anybody get there. The industry itself would not let itself get there. I can't imagine a plant management seeing 15 reds, 15 scrams and saying well gee, I still have seven to go.

MEMBER WALLIS: But that's for scrams.

But there are --

MEMBER ROSEN: That's this year. Another 22 next year.

CO-CHAIRMAN APOSTOLAKIS: Other than that, Mrs. Lincoln, I thought the show was good.

CO-CHAIRMAN SIEBER: Okay. Well, I'd like to hear a few words from everybody. Dr. Wallis.

MEMBER WALLIS: Well, I sort of agree with George. We've heard a lot of detail which I found very, very interesting. What this committee has to do is abstract from that a few things which are important where can influence, and George has picked out ones. I don't have anything to add to those. I think they're good items for the rest of the committee to think about and take a position on.

CO-CHAIRMAN SIEBER: Dr. Ford.

MEMBER FORD: Yes. I echo what has been said. On Davis-Besse, when this came up in discussion, you mentioned well, we didn't catch that. It was a green because we didn't have the ROP process being exercised for long enough. Is there any way of going back retrospectively to see if you would have predicted there was a performance issue at Davis-Besse?

MR. COE: I suppose that's possible, but

it would involve going back through a number of inspection reports prior to the implementation of ROP and trying to cast them into a different light in terms of the processes that we have today.

MEMBER FORD: Because it strikes your first bullet, when you say ROP is working effectively, and George very appropriately said that Davis-Besse said that you're not working effectively.

CO-CHAIRMAN SIEBER: Well, it depends on what you think working effectively means. I think that's a good choice of words because we don't know what it means. I don't think anybody has advertised ROP as being a predictor of anything. And, therefore, it's not a leading indicator, it's not a predictor. And you can't go back in Davis-Besse because the issues of interest occurred before ROP and the new system were in place. I think --

MEMBER FORD: My question, Jack, I said would such information be available so you could do a retroactive assessment?

MR. COE: I would have to think about that. I think the point that what is effective is a very good one, because from one perspective you could say that the self-revealing event or condition at Davis-Besse has been handled with a defined -- with

the defined process that's defined within the reactor oversight program, and that we have a special category for plants such as that that we're exercising now for Davis-Besse.

In other words, all of the tools that we have that have been utilized at other plants that have self-revealing conditions of significance are available and are being utilized in the case of Davis-Besse. And the question about could we predict another Davis-Besse in the future is a good one, and it's one that we ask ourselves a lot. And it motivates us to examine the operating experience that we do have, and try to find better ways of focusing our program and our inspections to help us find those things before they do become significant.

CO-CHAIRMAN SIEBER: It would be interesting though if you did find some leading indicators, and what would you do with the information? Could you go to the licensee and say you really haven't done anything bad, but you're going to. Okay? And what part of Title 10 do you stand on when you do that?

CO-CHAIRMAN APOSTOLAKIS: I think in all fairness, we have to separate the issue of what to do in the future from the existing, a disagreement

between the ACRS and the staff, or the apparent disagreement.

CO-CHAIRMAN SIEBER: That's true.

MEMBER SHACK: We'll set up a pre-crime unit like "Minority Report".

CO-CHAIRMAN APOSTOLAKIS: No, I mean this is the research everybody should think about. Even this committee I don't think has performance indicators to indicate -- to recommend for this particular issue, so this is for the future, but I was referring to the past. But coming back to the working effectively, the staff itself on page 8 of this thing writes, "It is important to note that the intent of these defining principles of the ROP was to result in an oversight process that provides adequate margin in the assessment of licensee performance, so that appropriate licensee and NRC actions are taken before unacceptable performance occurs." From that point of view, Davis-Besse is a failure.

CO-CHAIRMAN SIEBER: That's true.

CO-CHAIRMAN APOSTOLAKIS: At least, you can't say it's working effectively. I mean, this is in black and white here, "before unacceptable performance occurs." And I think we all agree now that that was completely unacceptable what happened

there. And yet, our own process said green. I mean, that's what the Chairman says. I hadn't seen -- but Chairman Meserve in his talk says yes, the day before it was green.

MR. SULLIVAN: I'm hoping this perspective helps, but when we were developing ROP, it was recognized -- two points were recognized that I'd like you to consider in your deliberations. One was that there were certain obscure issues we would miss. For instance, the D.C. Cook engineering problem that took place, it was just a closing as ROP was being developed, would not have been revealed by ROP either. It wasn't revealed by the core program, and it wouldn't be revealed by ROP.

And that leads me to my second point. It's not that ROP was ever claimed to be the absolutely perfect oversight program. We only thought that it was head and shoulders above the old one.

CO-CHAIRMAN APOSTOLAKIS: And I fully agree with you.

MR. SULLIVAN: Okay.

CO-CHAIRMAN APOSTOLAKIS: My disagreement is in saying that it's working effectively. I think we should be humble and say we did a good job up until now, though there are some disagreements. Now we

learn from experience we have to do something, but obviously it's not working effectively.

CO-CHAIRMAN SIEBER: Why don't we move on with the comments? Do you have anything else?

MEMBER KRESS: Yeah, I agree to a large extent with George. In particular, I would like to see this be viewed as a performance system and divorce it from risk almost entirely. And I think that's one of the problems.

I'd also echo his view that multiple findings ought not to be determined whether they're independent or not. They ought to almost assume that the root cause is such that they're related to each other, and they ought to be taken as an aggregate. I shared a consistency concern.

The question of how you should set thresholds of performance, we have a mixture now of judgment based on expert opinion and experience, and trying to use PRAs. I think that is one of the big problems we have with that, is throwing in the mixture. That ought to be based on judgment, expert opinion, and maybe use a Bayesian technique to improve on it as you go along. I don't think you should use risk to set thresholds.

I do think we need a different set of

performance indicators to deal with the boundary issues, as George says, and I think we ought to give some thought as to what those ought to be in order to be leading performance indicators that would pick out a degraded barrier a lot earlier than before it reaches a Davis-Besse.

As far as getting rid of the red, I think I'd keep it, but I'd sure look at the threshold, and change the threshold to a value that's meaningful. I think the red has significance in terms of, you may reach a red sometimes if you've got an appropriate threshold for it. And I think I would think about keeping it, but changing the threshold to an appropriate level.

CO-CHAIRMAN APOSTOLAKIS: Well, then I wouldn't disagree with that.

MEMBER KRESS: Yeah. I think that's basically the only expansion on what George said earlier.

CO-CHAIRMAN SIEBER: Steve.

MEMBER ROSEN: Yeah, thank you. I would disagree, being a confirmed rationalist here, that we ought to throw out risk. I think it works very well in initiating events and mitigating systems area, and that's really where it was intended, in my mind, to

function. It's been applied imperfectly in some of the other areas, and that's where we get into the trouble, so I would go to that point, where it's keep it for initiating events and mitigating systems, and apply it much more gingerly in the other areas.

I also have another worry that comes out of a visit of the ACRS to region two the last time we were there. Region two was very hospitable, brought in a whole bunch of people to talk to us, including a number of the residents on the plants in region two and the senior reactor analyst. And one of the things that we heard, which I've been sort of mulling on since that time and worrying about, was the statement by some of the residents that it was very, very hard to fit into their schedule the defense of a finding.

In fact, if they made findings in their work, the ROP, be they white, yellow, or you know whatever, it turned into a major, major work load for them. Maybe that was just because it's new. I hope so, but if it's not, if the process is so intensive, work intensive for the residents, the law of unintended consequences got us again.

We set in place a system. We rely on the residents to do it, and in fact, because it's so punishing they stop finding the things, or reporting

the things we want them to report. Now I'm not making an accusation. I'm just repeating what some of them said about their work load. They didn't say they were doing that. They just said it seems like it's almost very difficult for us to make a finding and then defend it.

You guys who are managing the agency using this process to upgrade it. And I agree, it's better than it was before, better than the process we had before. You need to think about the work load you're putting on the residents, give them all the help you can.

CO-CHAIRMAN SIEBER: Thank you, Steve. Dr. Bonaca. MEMBER BONACA: I pretty much endorse the perspective that George presented. One thing that I want to say, however, is that first of all, the presentations were helpful because I think I understood a number of things and reflections that you had. But it seems to me that since you're agreeing that the process is not cast in concrete yet, and there are opportunities for refinement, I mean, you could be open to some of the suggestions we are making here, or some of the inconsistences, because I think we discussed them, and you recognized some of them in certain cases. And I understand that probably it is

going to be hard to go back and say yes, performance is the issue, and the threshold should be not risk-informed. I mean, the threshold should not be -- but still you should consider doing that.

Anyway, that's -- so I mean, in general, the comments we put together in the original letter are still there, and really we haven't got any closure on that. And, you know, I think, however, in the context of again, your openness to consider improvements and the possibility of doing so, you should really -- it's going to be difficult for us to answer that SRM for the commission and say that we have worked with the staff at improving the process, because really we haven't been able to do that right now. That's pretty much that.

CO-CHAIRMAN SIEBER: Mr. Leitch.

MEMBER LEITCH: Well, I think the presentations today have been helpful, and my understanding, at least, of the aims of the ROP. I do believe that, as Steve Rosen has mentioned, that I think the initiating events and mitigating systems should still be risk-based, and others performance-based. I think that there's good basis for doing that, and I think it works well.

I guess I have a concern though in a

different vein, and that is, how we find -- and although the process isn't primarily designed to be predictive, yet I think what it is intended -- it is stated that it is intended to head-off things before they become big consequences. And I really think what gives us really big issues in the industry are not the individual things that are revealed by the ROP, but rather some way those things are summed and unexpected consequence, a major consequence occurs.

We've all mentioned Davis-Besse. We mentioned the D.C. Cook engineering issues. I guess in my own experience, harkening back to the operators asleep at Peach Bottom. You look at individual things, and I don't know how you get to some of these underlying problems, cross-cutting issues unless you drill down into those cross-cutting issues. If we stop our look, if we say that well, you're not going to look down that far because that's really beyond our scope, or beyond our charter to look down into those safety culture issues, into those management issues, I don't know how we find those things, because I think those are the things that really cause the industry and the agency big problems.

It's not the individual pump failure or valve failure. We can deal with those things. It's

those major cross-cutting issues that are the areas where we have high vulnerability. And I think we have to be looking down deeper into the process.

I know that's supposedly beyond our charter at the moment. I think there are some significant performance indicators that could be developed, that would give us a clue. Maybe not all the answers, but give us a clue as to some very important safety culture issues. And if we're not drilling down to look at those, I just think we're missing a big opportunity there.

CO-CHAIRMAN SIEBER: Dr. Shack.

MEMBER SHACK: Well, I do want to congratulate you on the presentation. I found it very helpful in understanding much of your rationale for getting towards the SDP, and coming up with things. I guess I'm fairly comfortable, if not total agreement would be expected, with a mix of performance-based and risk-informed, you know. I'd stay away from risk-based. I really think the notion here is to evaluate performance. And if I can use risk-informed views to do that, that's fine. If I have to use performance-based that's fine. And consistency from that, I don't see any other way you can do it except from judgment and experience, so I expect we will be adjusting these

performances as we do go along.

I agree with Graham that, you know, we do want to get back to -- you know, I see all this effort on the system indicators. You know, I actually think somebody would be doing -- you know, to go back to the other kinds of programs, you know, the problem-solving programs. You know, the system indicators I don't think are where the problems are at. And it's really the other kinds of performance we have to think about measuring. And we need more effort focused on that than we do developing new, more global safety system indicators. It's the corrective action program that perhaps is really the heart of what we're trying to know, as to how -- you know, you can't possibly inspect everything. What you have to have confidence is that the licensee's corrective action program is finding and fixing things. And, you know, that's where I would be focusing my efforts to look at performance indicators and better performance measures, not on my system performance. We can do that with the PRAs.

The ones you have may not be perfect, but as far as I'm concerned, they're probably good enough until I can handle other more important things that I don't think are dealt with, as well. I'd sort of

argue for some sort of re-focusing of the effort in developing performance indicators, I would see as the kind of highest priority I would like to see in improving the ROP.

MEMBER BONACA: Looking at different areas.

MEMBER SHACK: Looking at different areas. Again, the corrective action program is really the --

CO-CHAIRMAN APOSTOLAKIS: The completeness issue. Are we really --

MEMBER SHACK: Well --

CO-CHAIRMAN APOSTOLAKIS: It is an issue of completeness. You cover these, the staff, the systems, the hardware, the staff. Now we realize there's a hole there.

MEMBER SHACK: I mean, I also understand this need to have an objective program, and that really is kind of -- you know, you want to bury down -

- you know, the deeper you burrow, you know, the harder it is perhaps to come up with objective measures, but that's really where we need to be working.

CO-CHAIRMAN SIEBER: Thank you. I would like at this time to thank our presenters from NRR.

And I think it was very helpful, and gave us some insights that, frankly, we didn't have before.

MEMBER SHACK: Can I ask one question sort of off --

CO-CHAIRMAN SIEBER: Sure.

MEMBER SHACK: That's on the workbooks. You know, the other thing that we heard from the people, you know, using the workbooks and the SDP process was a bear. Do you think -- do you see ways to improve that?

MR. COE: Yes. There are ways to improve that. We have a task group that has just reported out, and has made some recommendations, and we're dealing with those now. And I believe that our ultimate objective is to improve the user-friendliness, if you will, of these processes. But in the very same breath, I will also acknowledge that using probabilistic tools in the program as intrinsically as we've made them a part of our program requires an additional intellectual effort. And we have to stand up to that and say we're willing to do that, and we'll make that as easy as it can be, as predictable, as scrutable, as understandable. But there's no question, and should be no question in anybody's mind that that is something we haven't done

in the past, and that we're expecting more not only of our inspectors, but of our staff management decision makers who make decisions at these SERP panels. So the answer is yes, but there will always be this need to make a greater effort to understand the tools that we're using because of their inherent complexity. There's no way around that, and so we have to acknowledge that.

Mr. Chairman, I have found this all very useful discussion. I'm pleased to hear that in some ways I think we satisfied the need to help give you a better understanding of our program. I would offer that the earlier letter that you sent us, although we may disagree that we need to redefine the theoretical basis for the program across all the cornerstones, we did agree that we need to be very much more clear about how we did design the program and its basis. From that standpoint, we believe that the scrutability of the program is our objective. And although we might debate the merits of one basis or another, what's important to us is the basis is clear. It's written down, and then we can debate something that hopefully is understood, and we can evaluate the different perspectives.

CO-CHAIRMAN SIEBER: Well, clarity is

important. Consistency is important, that we're in a position where you folks have a pretty well developed program, and we have a pretty well developed position, and somehow or other we've got to reconcile.

MR. COE: I understand. I was only reacting to Dr. Apostolakis' comment that maybe we didn't agree with anything in your previous letter. We did agree on that --

CO-CHAIRMAN SIEBER: Well, that would be a summary response.

MR. COE: So we thank you.

CO-CHAIRMAN SIEBER: I would like to take us off the record at this point.

(Off the record 4:34 p.m.)