Official Transcript of Proceedings

NUCLEAR REGULATORY COMMISSION

Title: Advisory Committee on Reactor Safeguards Human Factors and Reliability & Probabilistic Risk Assessment Subcommittees Meeting

Docket Number: (not applicable)

Location: Rockville, Maryland

Date: Thursday, December 15, 2005

Work Order No.: NRC-775 Pages 1-368

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

+ + + + +

ADVISORY COMMITTEE ON REACTOR SAFEGUARDS (ACRS)

HUMAN FACTORS AND RELIABILITY & PROBABILISTIC RISK ASSESSMENT SUBCOMMITTEE MEETING

+ + + + +

THURSDAY,

DECEMBER 15, 2005

+ + + + +

The meeting was convened in Room T-2B3 of Two White Flint North, 11545 Rockville Pike, Rockville, Maryland, at 8:30 a.m.

MEMBERS PRESENT:

GEORGE E. APOSTOLAKIS

MARIO V. BONACA

THOMAS S. KRESS

ACRS STAFF PRESENT:

ERIC A. THORNSBURY   ACRS Staff

ALSO PRESENT:

FRANK RAHN

ZOUHAIR ELAWAR

JEFF JULIUS

GARETH PERRY

JIMY YEROKUN

ERASMIA LOIS

JOHN FORESTER

ALAN KOLACZKOWSKI

SUSAN COOPER

MICHAEL CHEOK

DAVID GERTMAN

ANDREAS BYE

PER OLVIND BRAARUD

A-G-E-N-D-A

P-R-O-C-E-E-D-I-N-G-S

8:28 a.m.

CHAIRMAN APOSTOLAKIS: The meeting will now come to order.

This is a of the Advisory Committee on Reactor Safeguards Joint Subcommittees on Human Factors and Reliability and Probabilistic Risk Assessment. I'm George Apostolakis, Chairman of the Subcommittee of the Reliability and Probabilistic Risk Assessment Subcommittee.

Members in attendance are Mario Bonaca, Chair of the Human Factors Subcommittee and Tom Kress.

The purpose of this meeting is to review the status of the Agency's current research on human reliability analysis.

The Subcommittee will gather information, analyze relevant issues and facts and formulate proposed positions and actions as appropriate for deliberation by the full Committee.

Eric Thornsbury is the designated federal official for this meeting.

The rules for participation in today's meeting have been announced as part of the notice of this meeting previously published in the *Federal Register* on November 28, 20005.

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 first identify themselves and speak with sufficient clarity and volume so that they can be readily heard.

We have received no written comments or requests for time to make oral statements from members of the public regarding today's meeting.

We now proceed with the meeting and I call upon Dr. Frank Rahn of EPRI to begin the presentations.

Frank?

DR. RAHN: Yes. Thank you, Mr. Chairman, members of the Committee.

First of all, thank you for the invitation to appear before you and tell you a little bit about the program we have EPRI, in particular about the product for HRA, which we call the HRA Calculator.

Briefly an overview. We have three speakers with us today; myself, Dr. Zouhair Elawar from Arizona Public Service and Jeff Julius from Scientech.

This is a brief overview of what we intend to tell you. We have being passed out, I believe, copies of the presentation we have. And, of course, we will address the presentation to answer your questions as we go.

So just quickly, I think most of you know us but for those who don't, first I'll introduce myself. I've been with EPRI for 31 years. I'm manager of many of the risk and safety code applications at EPRI. And just a brief placing in some of my background.

We also have with us Dr. Zouhair Elawar from Arizona Public Service at Palo Verde Nuclear Generating Station.

Zouhair also has an impressive background. And I might mention, and he probably would be too modest to mention it if he did, but he's about to receive an industry award for the work he's doing on the HRA Calculator and the HRA users group.

And then lastly, Jeff Julius who, again, has very long experience, over 25 years in the nuclear business, many years doing HRA. Here is his critical information.

So you can see that between the three of us we probably represent 75 or 80 years of experience. That's kind of scary.

In any case, just a little overview of how the HRA Calculator project is working. First of all, EPRI manages the project on behalf of the industry and its members. EPRI has formed what we call an HRS users group whose purpose is to provide the guidance and resources to EPRI to develop the tools to guide us in our priorities and help us in terms of our quality assurance, beta testing, etcetera, prior to the release of the software.

Scientech is actually a contractor to EPRI, but functions to do the main development work, the maintenance, the QA testing, the training. This is directly funded work and, as you noticed from the first slide, that I have responsibilities with other of the EPRI projects. We do do jointly funded work, as an example, with the Risk and Liability User groups, since this is obviously an area of some interest in the to the PRA community. We have joint programs, joint training, etcetera and so on. And we try and coordinate all our efforts with other industry efforts such as our advisory committees with EPRI, the NEI, Nuclear Energy Institute here in Washington, various owners groups such as WOG and so on, BWR owner's group. And we have a number of international participants in the program also.

We will expand as we go along into some of these relationships.

Just a little bit of background. EPRI has been involved in HRA for a number of years. Many of you are familiar with and some of you have actually participated in some of these programs. The earliest work goes back about, like I say, 20 odd years. The first one was SHARP, which stands for Systematic Human Reliability Procedures in 1984.

We developed the HCR method, human cognitive reliability method in '84 also.

We're active in ORE and OPRAs, which are the operator reliability experiments and revised SHARP into SHARP1, and that was published. That was kind of precursor work to what we've been doing with the HRA Calculator.

At this point I'd like to introduce you to Zouhair. You already have his file statistics.

DR. ELAWAR My name is Zouhair Elawar. I work at the Palo Verde Nuclear Power Plant. And for the last ten years or so, the HRA work was my primary responsibilities.

The HRA Calculator group was formed about five years ago. So in my line of work I spend the first five years without the Calculator.

As I say, during the first five years, I spent the first two years doing HRAs about a couple of hundreds of them. And I have quickly realized that there is what is called analyst factor in doing HRAs. I have here a list of subtests that go into each HRA. And in each one of those items really you put the analyst factor as to how you will factor this into your HRA quantification, it has some subjective type judgments.

So which method you use or do you factor in alarms, accessibility, training, how do you factor the stress levels of operators? As you see all of those, you know, add a lot to the uncertainty in the HRA, which by itself have its own uncertainty from various NUREGs that we refer to get the values for operator errors in it.

Like I will mention, for example, like NUREG-1278, some people were using it as mean values, others were using it as median values. So there is a lot of uncertainty from the analyst factor in it.

So in the year about 2000 me and my peers realized that we need to form a group to come to the consensus in organized manner as to how to do this work.

Let me point out that the results used to vary widely between for HRAs from similar plans or even HRAs within the same plan; if you do the work over a period of about two years, you were in some mindset early on, you may have a different mindset a year and a half later. So I used to spend a lot of time doing consistency checks as to how did I resolve this issue six months ago, how am I resolving it today. So this was one of the main reasons why we thought we needed to have an industry group and form the HRA Calculator to come to convert to same methods with some consistency in it.

Later during our work we came to realize that we need also to form our Calculator to mirror ASME's HRA standard because we were getting a lot of peer review comments on HRAs.

I have to say that at this time because of the MISPI requirement all open comments on HRAs must have been resolved using the HRA Calculator.

CHAIRMAN APOSTOLAKIS: What did you just say? Say it again, please?

DR. ELAWAR The peer review comments on HRAs need to be resolved for a PRA model to be ready for MISPI applications. Any plan that have resolved those comments using the HRA Calculator, is considered.

I need to go back. Did i miss something

here?

CHAIRMAN APOSTOLAKIS: Let's go back to the slide, previous slide.

DR. ELAWAR Did I go back? Is this the one you wanted?

DR. ELAWAR Okay.

CHAIRMAN APOSTOLAKIS: Thirteen. Slide 13.

DR. ELAWAR Okay. Here it is.

CHAIRMAN APOSTOLAKIS: Have you tested your first bullet? Have you had different people using the same HRA method in obtaining comparable results?

DR. ELAWAR The testing is not formal testing, but we meet each year and we report among peers. I believe we are practically there. I mean, it's impossible to have it accurate in each application.

CHAIRMAN APOSTOLAKIS: Why is it impossible? Why can't you tell two different groups use the Calculator for the same sequence and compare the results? It can't be that difficult?

DR. ELAWAR I guess, yes, that's possible for one or two applications. When we are talking about a couple of hundred HRAs in each PRA model and the HRA Calculator when you go and start with it, you have to respond to scores of questions. You'll always have somebody really making a different judgment on one of the questions.

CHAIRMAN APOSTOLAKIS: Do two first, then worry about the 200.

DR. RAHN: I think if I might, Mr. Chairman, Frank Rahn.

The main testing really is coming through the peer review process. As Zouhair had mentioned, there has been extensive, I think as everybody's aware, peer review throughout the industry, the HRA. I think the peer review teams have been finding the consistency of the results between the plants that have been using the HRA Calculator.

CHAIRMAN APOSTOLAKIS: Do you have any hard numbers to show us, Frank?

DR. RAHN: We have an informal report on that.

CHAIRMAN APOSTOLAKIS: Okay.

DR. ELAWAR We can leave it as an open task and actually respond to you in some email in the near future.

CHAIRMAN APOSTOLAKIS: Okay. That will be fine.

DR. ELAWAR Yes. We can do that. That's

really simple. But I don't believe it was formally done, but nonetheless, you know, I have used it so many times. If I use it on one item and I use on something similar a month later, if I compare the results I say yes, great, they are consistent.

CHAIRMAN APOSTOLAKIS: Are you familiar with benchmark exercise that the European Union did about 15 or 20 years ago? it's a very disturbing figure that they show in a paper that was presented, I believe, in PSA-89. And we have to put that to rest at some time. We can't just ignore it.

What they did was they had the representatives from each countries of the Union plus the United States analyze the same sequence at a German plant. And they found that there was wide variability among teams using the same method, okay? And the same team using different methods.

At some point we have to do something about it. We have to demonstrate that the year of 2005 these things are not expected to happen again. So that's why your first bullet is of interest to me.

I suggest that you go and read that paper. It is only six pages and it reports on the results. And I know that everybody complains that this is very old and I keep bringing it up. But somehow, you know, we have to take care of it.

DR. ELAWAR Mr. Chairman, our own work before the Calculator was also pointing in that direction.

CHAIRMAN APOSTOLAKIS: Good.

DR. ELAWAR That's the main reason for our formation of the users group for HRA Calculator.

CHAIRMAN APOSTOLAKIS: Well, then we agree.

MEMBER BONACA: Just for example, you have a list of analyst factors.

DR. ELAWAR Yes.

MEMBER BONACA: Each one of them will have very subjective judgments. Now what have you done to make sure there is a common understanding of what, for example, operator stress level assignment is?

DR. ELAWAR We have now a clear guideline, I hope you will hear more from Jeff after me on this. We have a clear guideline now. You are in the Calculator, and you say okay now I have to enter a stress factor.

MEMBER BONACA: Okay.

DR. ELAWAR I click on help and all this appears, it comes in front of me, giving me a clear guideline. No vague guideline.

MEMBER BONACA: Okay. And this workshops that you say that you have among practitioners, you discuss how to interpret this guideline?

DR. ELAWAR Yes, we do. Let me say if I would say as to how we more or less eliminated that or diminished it.

If I go to start a new analysis, I don't go to my computer and start to work on it on the Calculator. Far from it. I have to go and prepare a whole, perhaps sometimes one week of leg work. I have in front of me a list, scores of questions, that I'm confident I will not miss anything in it if I am ready to answer them all accurately.

So, I go and do a week of leg work to be ready to go to my terminal and start to respond to those questions that are given to me in the guideline. And that is a key reason why I think that the analyst factor have been largely -- in fact, I believe, and I know as my peers too believe, that the uncertainty at this time using the Calculator, the uncertainty in the HRAs entered in the PRA model is very much comparable to other parameters, failure rates or initiating events that we put in the PRA model as well. I do not believe that we have more uncertainty from the HRAs.

And another point that I may make here up on my slide, through my peer review groups I participated in, the first questions that we go through are planned and want to examine the input of HRAs, we go to their model and answer all HRAs as true and we observe how a core damage frequency will change from, say, let's being 1 MLS 5 to becoming a 2 or a 3. Then we'll say, hey, we believe your HRAs are taking-- occupying the right place in your model.

CHAIRMAN APOSTOLAKIS: Two or three what?

DR. ELAWAR Two or three per year. If you go --

CHAIRMAN APOSTOLAKIS: Period?

DR. ELAWAR Yes. That's assumed the operator failed in every aspect.

CHAIRMAN APOSTOLAKIS: Presumably, you will not be able to see the second one, right?

DR. ELAWAR I agree with you. Until the frequency will -- if I go to a peer review and I see-- I put the HRAs, all of them, fail and I see the CDF remaining zero 0.1 or becoming 200, I wouldn't say your HRAs have something wrong in them.

CHAIRMAN APOSTOLAKIS: Now both you and Frank, I believe, mentioned the peer reviews. Can you give us some idea who the peer reviewers are? Not names. I mean --

DR. ELAWAR They usually are about ten 12 engineers, PRA engineers with various disciplines within the PRA.

Like when I go on those groups, they tell me you review the HRAs and you review the initiating events. I have more inclination to that area.

CHAIRMAN APOSTOLAKIS: Is that part of NEI peer review process? Is that what you're referring to?

DR. ELAWAR Is it part of NEI? Yes, perhaps. In fact, at this time the preparation of the PRA models to become acceptable for MISPI applications, all plants must close their peer review comments. And many plants have been reviewed prior to the Calculator being in effect, and they had HRA comments. I don't know that for a fact, but I assume they will meet their deadline and resolve those problems using the Calculator.

CHAIRMAN APOSTOLAKIS: Is anybody on the peer review team who is familiar with the various models of people who have proposed internationally who is familiar with some of the psychological literature, or are they all engineers?

DR. ELAWAR They are all engineers.

CHAIRMAN APOSTOLAKIS: All engineers.

DR. ELAWAR All experienced PRA engineers.

CHAIRMAN APOSTOLAKIS: Okay.

DR. RAHN: And, Mr. Chairman, we'll come back to that question later in the presentation.

I'd just like to make a comment explaining Mario's observation of that training. One of the key things that we've been doing in the users group is holding usually at three training sessions a year where we have on average about 20 folks attending each one of those. We are starting to come to a consistent understanding within the community and building up a cadre of people who have similar trainings so that the communication and the models that are being used are consistent between plants.

I think that's a rather key point.

CHAIRMAN APOSTOLAKIS: Are you coming back to the training issue later?

DR. RAHN: Yes, we will talk about training.

CHAIRMAN APOSTOLAKIS: Okay. Let's move on to slide 14.

MEMBER BONACA: This is great. And the only thoughts I still have on this is that, of course, once you have consistency of interprotection doesn't mean that is providing the answer. I am is there anything that you do to verify, for example, against simulator exercises and so on? You don't have to answer now, but at some point in the presentation there will be some discussion of it.

DR. ELAWAR Actually, operator review and simulator exercises are part of each HRA analysis. When I do one HRA, I prepare a list of my assumptions and responses to questions. I document them and before I --

MEMBER BONACA: So you will discuss later at some point?

DR. ELAWAR Yes. We will go to the operators' training and operators. And we see we don't ask them to give us answers, because usually they are optimistic than they ought to be on this issue. I go and say, look, I am making those assumptions, it's in the procedure I say that the operator is going to do this and this and this. And I think I'm assuming it will take him ten minutes to do this work. The operators' training or the senior reactor operator will say yes or will correct me if I'm wrong.

So, in fact, the operator involvement is very, very heavy in HRAs.

MEMBER BONACA: Okay. All right. Thank you.

DR. ELAWAR And that's if I'm in a peer review of work and I will see a documentation of operator involvement, I will put as a type A comment you have to take HRAs and have operators review them and comment on them, and agree to them sort of back there. There were many comments of that nature.

MEMBER BONACA: Okay. Thank you.

DR. ELAWAR Any questions over here? Did I miss anything here?

I guess I will have to say finally that I am very confident with the HRA Calculator applications as being so comprehensive that it has in it, it would alert you to so many questions and given you guideline to respond into them that what I believe used to be a heavy analyst factor --

CHAIRMAN APOSTOLAKIS: Can you give us an example of a question or two?

DR. ELAWAR On the Calculator?

CHAIRMAN APOSTOLAKIS: Yes.

DR. ELAWAR I think you are going to see most of them presented on slides today.

CHAIRMAN APOSTOLAKIS: Okay. Fine. Fine.

Now go back please.

DR. ELAWAR I apologize for this. I'm not clear. Which slide number do you want to see?

CHAIRMAN APOSTOLAKIS: I don't know. What was it? Fifteen.

MEMBER BONACA: Fifteen, I think.

DR. ELAWAR This is simply --

CHAIRMAN APOSTOLAKIS: Yes. If I wanted to access these websites, I have access to the first one, right?

DR. ELAWAR Yes. See, we have --

CHAIRMAN APOSTOLAKIS: Our membership--

DR. RAHN: Yes, it's both a public and private website. The first one is the public website where anybody, members of the public can get international --

DR. ELAWAR We have 22 user groups participating.

CHAIRMAN APOSTOLAKIS: I'm asking about me. Which ones of these can I access?

DR. ELAWAR You can go to the --

DR. RAHN: The top one is --

DR. ELAWAR -- public website. Because not all reviews are participated and paying for it. So there are some activities that cannot access the Calculator per say itself.

DR. RAHN: But most of the information in the users group is in the public website. The next bullet it says what website, that's mainly for downloading of software products which are supported by the users group.

CHAIRMAN APOSTOLAKIS: But if I wanted to understand what assumptions you are making and how you are producing the results, would the public website be sufficient for me?

DR. ELAWAR Probably not. I think you have to review. I can personally send to you a sample HRA from my files --

CHAIRMAN APOSTOLAKIS: Well, send it to Mr. Thornsbury.

DR. ELAWAR Okay. I can do that.

CHAIRMAN APOSTOLAKIS: He is a trustworthy guy.

DR. ELAWAR In the documentation, actually if I press my documentation button, it will give you few pages of everything you have assumed and where you quantified it from. In other words, a technical reviewer looking at the documentation put out on the HRA Calculator it is such that he doesn't have to go back to the preparer and ask questions.

CHAIRMAN APOSTOLAKIS: Are you familiar

with the work that this Agency has been doing on human reliability the last 15/20 years?

DR. ELAWAR I am very familiar with NUREG-1792 was put out as the good practice. We think it's a great, great document.

CHAIRMAN APOSTOLAKIS: About some of the other work they have done? I mean, ATHEANA, are you familiar with ATHEANA?

DR. ELAWAR I am familiar with ATHENA, familiar -- oh, yes. We use NUREG 1278 extensively for our quantification.

CHAIRMAN APOSTOLAKIS: So there is a number of models out there, as I am sure you are aware of, right?

DR. ELAWAR Yes. Yes, I am.

CHAIRMAN APOSTOLAKIS: SPAR-H, are you familiar with SPAR-H?

DR. ELAWAR I'm very familiar with SPAR-H. Yes. I mean this is --

CHAIRMAN APOSTOLAKIS: If somebody looks at these models, one gets the impression that most likely if I use two of these, I'll get two different answers, right?

DR. ELAWAR Well, two different answers is a relative term. Obviously, you would not expect the exact same answer --

CHAIRMAN APOSTOLAKIS: They're not the same, they're different, right? What do you mean it's a "relative term?" There are two different answers. SPAR-H says. you know, the nominal error rate for errors of diagnoses is about 1- to the minus 2, I think. And then they adjust it. Other methods may give something else.

My question is, and I think this is a realistic you have the current state of the art.

DR. ELAWAR Yes, yes.

CHAIRMAN APOSTOLAKIS: I'm not saying it to blame anybody. Is the EPRI Calculator eliminating these differences?

DR. ELAWAR Those differences as I see them now, they are within the error factor for that answer you are getting.

CHAIRMAN APOSTOLAKIS: Yes.

DR. ELAWAR And that's one thing. And the other thing you have to look at it in the aggregate as to if I am doing 100 HRAs and the other person doing the same 100, I may be higher on one or two here and lower on one or two there and vice versa. But in the aggregate we should be really very consistent.

CHAIRMAN APOSTOLAKIS: There is a

difference between "we should be" and --

DR. ELAWAR We are.

CHAIRMAN APOSTOLAKIS: -- "we are." We are?

DR. ELAWAR No. I'm saying we are.

CHAIRMAN APOSTOLAKIS: And do you have any evidence of that?

DR. ELAWAR Well, really, talking with peers and remembering myself as to what I did six months what I do now, and in meetings how people stand up and speak of it as it being to that degree of accuracy. But it's not --

CHAIRMAN APOSTOLAKIS: Are you saying that it doesn't matter which model I use if I --

DR. ELAWAR No, I'm not saying that.

CHAIRMAN APOSTOLAKIS: -- put uncertainty bounds, I more or less find the same range?

DR. ELAWAR Not quite so. I think there are models of more importance, and I have to say that a great majority of our users rely on the third quantification model. And those who are using that third model, like I am at my plant, they will be largely consistent.

If I have an HRA with a result of 2a-3, somebody else may have a 2.1a-3 and another person might have a 1.8a-3 with an error factor of say, 5. I will still view those as being consistent.

CHAIRMAN APOSTOLAKIS: I would, too.

DR. ELAWAR Yes.

CHAIRMAN APOSTOLAKIS: What worries me is if one guy says ten to the minus 5.

DR. ELAWAR If I one guy say that, the peer review will likely catch it. And I believe that is extremely rare for this issue. This extreme difference is very unlikely with qualified people using.

Let me also add one more idea, an HRA practitioner using the Calculator is not somebody who is simply being trained how to use it. The person has to be a PRA qualified person and then have to go through 3 or 4 days of training.

CHAIRMAN APOSTOLAKIS: Well, what does that mean? What does that mean PRA qualified? I mean, there --

DR. ELAWAR He has to know how to put fault trees, event trees, how the water systems -- he has to know --

CHAIRMAN APOSTOLAKIS: Has to have done it before, you say?

DR. ELAWAR Yes. He has to know how to do

PRAs. Only after you are a qualified PRA engineers you can go and be trained to do HRAs.

CHAIRMAN APOSTOLAKIS: Okay.

DR. ELAWAR I do not expect to see such large differences --

CHAIRMAN APOSTOLAKIS: You are giving us a more optimistic view than I have. But I am willing to be convinced.

DR. ELAWAR I am saying my bottom line is the uncertainty in the HRAs with the Calculator are comparable to the uncertainty of our parameters such as component failures and initiating event frequencies.

CHAIRMAN APOSTOLAKIS: But there's not a big difference there. I mean, for component failures at least you have plant specific data for most of it so you can update your distribution and feel more comfortable with it --

DR. ELAWAR Yes, you still have to put--

CHAIRMAN APOSTOLAKIS: With HRA it's a little the judgment of people, isn't it? I mean, you can't update any --

DR. ELAWAR Well, let's see, if you look at NUREG-1278, it's a 1,000 page document specific to nuclear power plant applications with so many expensive tables and information in it, I mean that's what we go -- usually we go by in quantifications.

CHAIRMAN APOSTOLAKIS: Great. Thank you. You have anything else?

DR. ELAWAR I'm ready to answer questions.

CHAIRMAN APOSTOLAKIS: Okay. Is there another presentation from EPRI?

DR. RAHN: Yes.

CHAIRMAN APOSTOLAKIS: Let's go on.

DR. RAHN: Frank Rahn again. To follow on with some of the comments that Zouhair has just made. I'll expand a little bit on our technical approach.

We have a specific mission when we started this five years ago, and that is first of all, we wanted to ensure that we would have a software tool that would meet the regulatory and safety analysis needs of our members. And we needed tools that we could use essentially right away. We didn't have 5/10/20 years to do large research programs because it was obvious that the need was critical.

We wanted to have defensible and reproducible reports. We wanted to be able to automatically produce reports that would have common formats or that when the reviewers would come in, they would have an opportunity to look at something, a format that they would be familiar with and they wouldn't spend a lot of time trying to decipher what was done, why it was done and so on.

so as a result we turned to the methods that really had been widely used up to that time, and they're still widely used now. So we would have an industry-wide understanding of what was going on.

We had a couple of essentially criteria for what we were doing. We wanted to have tools that would be traceable. We wanted to have tools that would be defendable. We wanted to have tools that would be consistent.

We recognized that whatever we picked there would be some things that were on the positive side and some things that were less well understood, but at least we wanted tools that we understood both the strengths and the weaknesses of those tools such that we could then use that as a basis for moving forward.

So in addition to that we developed manuals and help to work with our software. We wanted to promote consistency. Like I said, we have usually about three per year training sessions, well attended. We usually get about 20 to 30 folks that come. We've been doing this for three, four, five years now so you can see we're starting to build up a cadre of industry personnel that are thinking alike, using the same types of assumptions. We document those assumptions. it doesn't mean necessarily that we always get the right answer, but at least we understand what we're dealing with.

Of course, we want to map with the ASME PRA standard, which is recently out. And we do that directly either through something called EPSA, which is a software tool which essentially allows utilities to document criteria by criteria in the standard and essentially state to what level that they meet the standard and where the shortcomings are and where the assumptions are.

There's also something we're working on now which is not ready yet, but we will have shortly called Document Assistant, which again is where it's permanently documenting the results such that they don't get filed away in a cabinet someplace and five years from now nobody can find them anymore.

And then lastly, we focus mainly as the standard has on the level 1 PRA or PSA, and we're building the foundation for future, certainly with the SDP process, we're expanding out into the fire and flood area, shutdown area. So these are still areas of development, but we are starting to make progress there also.

We work with universities. Most recently with Texas A&M so if you are familiar with Bill Virgil there. We've had recently one or two students producing master's thesis using the Calculator and producing a report. We hope to expand that in the future to other universities. We do make our software available to universities, essentially at a nominal cost for their use and for their training purposes.

We use the user group now is a focal point, a way if you will, mustering industry resources to essentially work interactively with NRC. Occasionally we get requests from NRC to review various of their documents. So EPRI works with the users group to coordinate the responses to those documents, uses those documents as a way of comparing what we're doing with what NRC is doing and some of the things we've commented on the NRC Good Practices, the SPAR-H models, the HERA, the Human Events Repository.

We also have international members. That allows user groups to have a wider, if you will, view of the world, what's going on internationally.

There's, as you know, programs going on particularly in Europe, a number of places there, Germany, Finland, etcetera have been very active in this area. They have been producing new ways of doing things.

We test them occasionally. One of our international partners was EdF in France. We explored a method that they're developing right now called MERMOS. And we will continue to do so. But right now, unless a methodology has been well tested and is out there for a number of years that we can use with some confidence, we are I might say a little bit on the slow side to adopting it. Because we want to use well tested methods and we understand that in the future there may be better ways of doing things, but until we understand all the ups and downs of these new methods we're probably not ready to implement them.

CHAIRMAN APOSTOLAKIS: Can you tell us a few words about what you actually said on these documents? I mean, you told us that you reviewed them. What do you think of the Good Practices documents, SPAR-H --

DR. RAHN: Well, I think both of those are certainly the Good Practices, a good step forward. And you know, we've taken some of the -- well, actually most of the suggestions there and we incorporate them in the way we do our Good Practices.

I think we had a few comments back or we had a few suggestions. But by in large, I don't think there are any major disagreements between what NRC was thinking and what we were thinking.

In fact we have incorporated and you will hear in the next presentation how we incorporate SPAR-H into our methodology. So we have high regard for the things that NRC is doing and has done.

CHAIRMAN APOSTOLAKIS: But if you go --

DR. ELAWAR If I may add, SPAR-H is not for use by the industry, it's just for comparison purposes. Whatever you are using, you say well if the NRC is using with SPAR-H, what do they get compared with what I do. It's not meant to be used by the industry.

CHAIRMAN APOSTOLAKIS: Why not?

DR. ELAWAR Well, some people may decide to use it, but I don't know of anybody that uses it--

CHAIRMAN APOSTOLAKIS: You said "it's not meant." Do you think the authors of the report did not want other people to use it?

DR. ELAWAR Well, see like other PRA models for various reasons are also with NRC in a simplified manner. It's not as detailed as we like to use the method. As far as I know, whether it's right or wrong, utilities are not using SPAR-H --

CHAIRMAN APOSTOLAKIS: Well, maybe it's because they're now.

DR. RAHN: Well, I think it's more, George, that you know NRC has developed an independent way of reviewing what industry is doing.

DR. ELAWAR Correct.

DR. RAHN: And if we're using the same tools, you really don't have your independent view, if you will. So we in the industry we like to compare against SPAR-H because if our answers are grossly different from what NRC would be getting, that's obviously a flag that we're on the wrong track.

CHAIRMAN APOSTOLAKIS: How about MERMOS, what do you guys think of that?

DR. RAHN: Well, MERMOS is a tool that's been developed at EdF, it's essentially the post-accident. Our view is that it's a technique under development and hasn't been used long enough at EdF or other utilities for us to adopt it at this time. And that's going to be said of a number of the other techniques.

We are interested in things that have been out there for a while and are well tested. And, again, they're not perfect but at least we will understand what the weaknesses are and where the strengths are, and that allows us to move forward with confidence.

So right now the models that we are using in the HRA Calculator, the THERP model, that obviously goes back a number of years and a NUREG report started it. I think that goes back about 1980 --

DR. ELAWAR 1983.

DR. RAHN: '83/'84, that time frame.

The ASEP model, again, another NRC NUREG on that, 4772. And those are for the pre-initiator. HRA for the post-initiator HRA we're using CBDTM, which is a caused based decision making model and in combination with THERP. We have the HCR/ORE/THERP models, the annunciator response model, a combination of the cause-based and the HCR/ORE. And that was in an EPRI report 100.259.

And then the THERP annunciator response model.

So we have a number of models that are built in --

CHAIRMAN APOSTOLAKIS: Doesn't the ASEP deal with post-initiator errors, too? I thought the ASEP did that?

DR. RAHN: Well, it does. But we are

using it primarily right now for the pre-initiator part of the --

CHAIRMAN APOSTOLAKIS: So the primary model for post-initiator is which one?

DR. RAHN: Is the cause-based decision, it's what most --

CHAIRMAN APOSTOLAKIS: CBDTM?

DR. RAHN: Right.

CHAIRMAN APOSTOLAKIS: Which includes HCR or is it different?

DR. RAHN: It's different. Jeff will explain in the following presentation the details of the various models.

What's new recently meaning in the last year? We have been concentrating on the following points trying to improve the software we have. Certainly the dependency analysis function where we are looking at how dependencies influence our answers.

We're looking at links between performance shaping and the quantification itself.

Certainly we are integrating with the ASME standards here. We've included the SPAR-H model and the next presentation, which Jeff Julius will give you some of the details on all of those.

MEMBER BONACA: The question I have is

clearly you made -- you know going back to 19. A selection of different models that exist already in the industry for different characterization; pre-initiators, HRA you have chosen certain models. You have chosen not to use SPAR-H, you have chosen not to use ATHEANA. So how do you go about making the selection of programs that you use now in the Calculator? Did you make some comparison?

DR. RAHN: Well, I must say we had, call it a fairly pragmatic approach in the sense that when we first started the project five years ago or so we looked at the types of things people were using. And for us, and as Zouhair explained, a lot of them were all over the map. So our first step was to build on that base and try and bring people together. So we tried to incorporate in the HRA Calculator the models that were being used in the industry and then start to move forward through a common model. So we started with a number as indicated by this slide of the commonly used methods. And we're starting to grow into a more common approach how to do HRA.

MEMBER BONACA: But you had to make yourself comfortable that in fact even if it goes unused by the Agency before was appropriate and adequate for the job to be done?

DR. RAHN: That's done.

MEMBER BONACA: And it wasn't missing certain elements. So you did also that kind of selection? I mean, it wasn't only based on --

DR. RAHN: Right. Exactly.

And with that, Jeff?

MR. JULIUS: Good morning. My name is Jeff Julius. I work for Scientech. I've been in the nuclear industry for 25 years, approximately 16 years working on human reliability and the last few years with EPRI.

And this portion of the presentation we'll describe the methods and the approach used in the Calculator. As you've heard from the preceding slides, the Calculator itself is primarily a tool and that there are other aspects that are involved with the HRA user group such as the guidelines and the training to promote the consistency and the standardization of the approach to HRA.

In general, the HRA Calculator technical approach, it follows the ASME and SHARP framework. The general process for identification, screening, the qualitative characterization and the quantification and dependency evaluation of the human failure events.

One of the things that is the key output of this process is both the qualitative insights as well as the quantification of the human error probability. Obviously if we had actuarial or historical data, we wouldn't need to develop some of these scheme of models, but unfortunately we don't. We don't have a lot of historical data for these types of events. So we break down and the Calculator approach has been to integrate and use previously developed research and models.

To answer one of your questions, this development process has pretty much gone along in parallel with SPAR and it was drawn from, you saw from the proceeding slide, NUREG-1278, the EPRI reports TR-100.259 which culminated, started with simulator experiments and then developed this cause-based decision tree approach. So we've kind of combined and packaged and integrated to allow the different selection methods as well as build on the lessons learned during those ten years from doing the different human reliabilities.

So we start with the input of the qualitative factors. And we promote consistency by standardizing the definition of the qualitative performance shaping factors.

CHAIRMAN APOSTOLAKIS: But let me

understand the second bullet. Allows for selection of methods. On what basis? I mean, what advice do you give to the user as to how the select the method?

MR. JULIUS: The advice that we give to the user is to start with the cause-based decision tree. For example, for the post-initiator events. Start with the cause-based decision tree method. In THERP the cause-based decision tree method, as you'll see, has a series of questions that are asked regarding the man machine interface in the cues and then the procedures. And that produces data, qualitative data and probability results. And then we look at that value and we look at the timing aspects. Human cogitative reliability method is better used for the short time frame scenario actions where the operator response is more time driven. The cause-based decision tree is given he's got plenty of time, what are the different factors.

CHAIRMAN APOSTOLAKIS: Well, let's talk about the HCR. As you know, some people are questioning the basic assumption of the log normal distribution there. There's a log normal distribution for time, it gives it a probability of not taking action, I think.

MR. JULIUS: In a sense, normalized, yes.

CHAIRMAN APOSTOLAKIS: Given a particular time. And people have questioned that. And I believe the new document from the NRC comparing with the Best Practices mentions that.

If I am a user and I go to the EPRI Calculator and I look at these models, is there anything under HCR that will tell me that some people might question this in the future? If you do this, you're taking a risk?

MR. JULIUS: No.

CHAIRMAN APOSTOLAKIS: Are you questioning the assumptions of the models?

MR. JULIUS: No, we have not questioned the assumption of the model. And in general, the human reliability area has been that anything you put down is subject to question in the future, whether it's the cause-based decision tree or the HCR.

CHAIRMAN APOSTOLAKIS: Some things are more questionable than others.

MR. JULIUS: Yes. But one of the points we do question and point out is because it uses this log normal and normalized -- the log normal approach to the time, is that the human error probabilities can drop off to very low values very quickly. So that, for example, if your timing window is 20 to 30 minutes and your median response time changes from 15 minutes to 10 minutes, that can produce two or three orders of magnitude difference. And the time window expands to 45 minutes or an hour, you can produce a 10 to the minus 14 or 10 to the minus 15 human error probability if you blindly apply the approach.

What the Calculator does then is to say, wait a minute, that's too below, below the minimum believable.

CHAIRMAN APOSTOLAKIS: Now your statement earlier that all HRA methods have questionable assumptions, are you saying then that all of them are equally valid or equally invalid? Are some methods that are better than others, perhaps? All of them are questionable, therefore I don't care about it?

DR. RAHN: This is Frank Rahn.

We have a rather different approach. We want to be able to document and record what we've done. Document our assumptions. So that if it turns out in the future that some efforts are proven to be much superior to the ones we're using, we'll be able to go back and understand where we need to make adjustments.

CHAIRMAN APOSTOLAKIS: I don't know how a method can be proven to be inadequate.

DR. RAHN: Well, as you point out, some might be maybe more adequate than others.

CHAIRMAN APOSTOLAKIS: But there is a tendency, I believe, in this field not just on your part but in general, people they feel they have to list a number of models. And they say well this and that and this and that, there's some discussion. But nobody is willing to say this is plain wrong or this is an assumption that has no basis on anything.

Now, you can't expect the PRA users to go so deeply and study ATHEANA, study CREAM, everything, and say my God, you know Nogel says this on page 232 in his book and I disagree with that. Somebody has to do that. And by saying, you know, we're only going to list models that have been used, I don't know how that helps anybody. I mean, you have to have some sort of evaluation there.

For example, coming back to the HCR, these median times, I think the recommendation is to actually do plant specific performance experiments and get it with operators. Now that's probably not an inexpensive effort. Are you saying anything about that there or are people going to use some sort of generic number or they will ask the operator what do you think and the operator will say 3 hours, and everything is fine?

DR. ELAWAR: If I may make a comment here?

CHAIRMAN APOSTOLAKIS: Of course you may.

DR. ELAWAR: The HRA Calculator is not the only source for somebody shopping for a method. When we start to do the work it is my plan before the HRA Calculator or somebody or two person spent weeks and weeks reviewing what's available until they have decided I am going to use this for this application and this for that application. So to answer your question, yes they do look in detail.

CHAIRMAN APOSTOLAKIS: No, they can't.

DR. ELAWAR: Not for each application. Like for example, I use THERP for quantification and I use it consistently. I don't go look for other methods if I've applied an answer here or there.

CHAIRMAN APOSTOLAKIS: Well, one of the precedents that this draft NUREG does is the comparison of HRA models with Good Practices document, is that it has usually half a page of commentary after each method. And it lists maybe advantages, disadvantages, what is questionable. It seems to me that something like that should be extremely useful to your users if after each method you put something like this or to say wait a minute, now if you use this method it contains this particular assumption which may be questioned in the future. And maybe you don't want to invest, you know, whatever it takes to do the HRA and then have somebody say well you don't believe it.

DR. ELAWAR: I believe --

CHAIRMAN APOSTOLAKIS: That is a great step forward, is it not?

DR. ELAWAR: In my report, although HRAs which is about 200 pages, the first 40 pages are dedicated to analysis of methods; how did I go about selecting what I want to use and it contains that information specifically as you have mentioned. And then--

CHAIRMAN APOSTOLAKIS: Well, that's good.

DR. ELAWAR: So in other words, there is really a long time spent in each comprehensive HRA report. It starts with the declaration of which methods I'm to use, which ones are available, which ones are better for what application, a declaration of principles sort of, and then the actual --

CHAIRMAN APOSTOLAKIS: What do you mean what methods are better for what application?

DR. ELAWAR: Like, for example, I said okay here I want to use three or four quantification and I have several pages describing myself as to why I made that decision. What I look at as well to come to this conclusion.

CHAIRMAN APOSTOLAKIS: Well, let me put it different. Okay. I do that. Then is it possible that there will be another, say, fact somewhere or accident sequence where you will advise me not to use THERP because of something else there?

DR. ELAWAR: If I knew of that, I will. I don't know that I know of that in terms of using THERP for quantification.

CHAIRMAN APOSTOLAKIS: But isn't it the case where a guy selects the method and then uses it everywhere? I mean, for post-initiator it may be different from pre-initiator. But if I decide to go with the decision tree, then all my post-initiator events will be done that way. I can't imagine that people say, hey, I'll do it 70 percent of the time.

DR. ELAWAR: Yes, that is logical.

CHAIRMAN APOSTOLAKIS: But there are these other things here that I have to do something else with.

DR. ELAWAR: Yes. Well, we try to --

MR. JULIUS: Well, a lot of them do.

CHAIRMAN APOSTOLAKIS: So you're saying

that you have actually evaluated -- I mean, have you seen this draft NUREG?

MR. JULIUS: No.

DR. ELAWAR: I have actually evaluated--

CHAIRMAN APOSTOLAKIS: You have to speak with sufficient clarity and volume.

DR. ELAWAR: I apologize.

I did actually evaluate, in other words I say in my report I have about 40 pages dedicated for the reader to know how did I go about selecting. It's not -- the Calculator is an abbreviation of that. It's just simply a reminder to the user, hey, this method is method for this or it is for that, but this is not really what the users have relied upon to come to a decision as to which method to use.

It is a detailed, up front evaluation that was done even before the calculation.

In my case I am confident that work --

CHAIRMAN APOSTOLAKIS: I mean, if you can give us examples. I mean, if you can send Eric here with documents --

DR. ELAWAR: I am permitted to do that. I will send them to Eric.

CHAIRMAN APOSTOLAKIS: That will be great. Because, you know, that will help everyone.

DR. ELAWAR: But for your information if you look in this report, you will not simply start with item number one here it is, that's the analysis. It will not start like that. It will start with detailed discussions about the principle, how do I look at methods, how am I going to deal with operators, what kind of assumptions I'm going to make. It's a declaration of principle. I will stick to it further on instead of I don't like the answer by this method, I'm going to look for a --

CHAIRMAN APOSTOLAKIS: But when you do that are you saying and this model appears to be the most compatible one with what I want? You're not saying that?

DR. ELAWAR: Well, I am saying that by--

CHAIRMAN APOSTOLAKIS: You're saying that? Okay.

DR. ELAWAR: I mean, not in the same words. But by saying I learned of those methods and I believe because this method have those characteristics, I'm using this third model for quantification.

CHAIRMAN APOSTOLAKIS: Okay.

DR. ELAWAR: With several pages describing it why I made that decision. Obviously, I would have preferred it over other available methods.

CHAIRMAN APOSTOLAKIS: Fine. If you do that, that's fine. Then we agree.

DR. ELAWAR: Yes.

CHAIRMAN APOSTOLAKIS: Don't be surprised and look at me that way. We can agree every now and then.

DR. ELAWAR: I appreciate that.

CHAIRMAN APOSTOLAKIS: You look so stunned.

DR. ELAWAR: I understand the PRA model is a docketed document. That's why, I mean, it's not available for NRC reviewers in details.

CHAIRMAN APOSTOLAKIS: Don't --

DR. ELAWAR: Well, I mean lack of --

CHAIRMAN APOSTOLAKIS: Don't tell me that. Okay. If you submit something to this Agency for review, an application, this Agency should have the right to review the model.

DR. ELAWAR: Well, nobody's doing that right. But the fact is --

CHAIRMAN APOSTOLAKIS: I understand they don't have the data that were developed during the ORE.

DR. ELAWAR: That's why --

CHAIRMAN APOSTOLAKIS: So anything that comes with HCR here should be rejected, in my view.

So let's go on.

MR. JULIUS: So what's the title of that NUREG? We are familiar with --

CHAIRMAN APOSTOLAKIS: Oh, that's a draft.

MR. JULIUS: That's right. And I don't believe we've seen that. We know that the NRC has got a series of --

CHAIRMAN APOSTOLAKIS: Well, are you here today?

MR. JULIUS: Yes.

CHAIRMAN APOSTOLAKIS: They're going to present it right after you.

MR. JULIUS: Okay. But you asked if we had seen it yet, and --

CHAIRMAN APOSTOLAKIS: No, that's fine. Yes, draft reports are not published, right? The report is not published.

DR. LOIS: (Off microphone).

CHAIRMAN APOSTOLAKIS: You are away from the microphone. So Dr. Lois just said that the report is not published.

MR. JULIUS: Okay.

CHAIRMAN APOSTOLAKIS: So we all agree

with you. Okay.

MR. JULIUS: All right. The bottom bullet here then. We promote consistency by standardizing the definition of the qualitative performance shaping factors. One of the things we saw between the different plants was that different definitions of the timing and the time windows.

Promote guidelines for the selection of performance shaping factor and characteristics.

CHAIRMAN APOSTOLAKIS: So you are giving definitions for the various PSFs, Jeff, is that what you're saying?

MR. JULIUS: Yes.

CHAIRMAN APOSTOLAKIS: Now you said something about timing. Is there any question there that people don't understand what we mean by it?

MR. JULIUS: There are some questions. For example, we had one of the human interactions I reviewed was a utility that said, hey, we've got a six hour time window for this action so the human error probability must be low, 10 to the minus 3, 10 to the minus 4. And then when you actually laid out the time window and followed the event tree it was one of these actions that it was restoration of emergency core cooling system after a station blackout. Well, the restoration on the event tree didn't start until the power we recovered at 4 hours into the event. And then the amount of time it took for the manipulation time, to get the breakers and get the support systems aligned that you could start the front line systems basically left out of that 5 or 6 hour time window a half hour or 45 minutes to complete the action. And they didn't account for this delay.

So the laying it out in a standardized framework with accounting for the delays and the manipulation and the time for the cognitive response gives a clearer timing and a consistent timing picture. And you'll see that in one of the graphics in the next slide.

CHAIRMAN APOSTOLAKIS: Okay.

MR. JULIUS: The other thing on the guidelines for some of the selection of the performance shaping factors. This has been a evolutionary approach. I think even in version 2 that was reviewed by -- the software that was reviewed and used in that draft NUREG we started out in version 1, you know, here's the model we have. We put it into some software so we can do quicker updates.

The version 2 came after ASME and ASME said well you need to look at these performance shaping factors. And some of them we hadn't looked at before. So we said okay, now the software forced you to look at it but there was a disconnect between the qualitative and the quantitative story.

And in this version 3 now we have a tighter connection. Okay. If the action is complex or if there is some negative performance shaping factors, that should drive an increase for example in the stress.

CHAIRMAN APOSTOLAKIS: So do you have a list of performance shaping factors and then some advice which ones might be important to the particular event?

MR. JULIUS: Yes, we have a list of performance shaping factors. And we actually shared that with the NRC Research when they were developing the HERA database so we could make sure that we -- and we've also compared them with SPAR to see the consistency and the general performances shapes and factors.

CHAIRMAN APOSTOLAKIS: And what kind of guidelines do you have there? How do people select the PSF?

MR. JULIUS: Well, you'll see here in a subsequent slide.

CHAIRMAN APOSTOLAKIS: Okay.

MR. JULIUS: Let me get to that.

CHAIRMAN APOSTOLAKIS: All right. Let's move on.

MR. JULIUS: Okay. This is, again, the different types of models and the features.

CHAIRMAN APOSTOLAKIS: Yes.

MR. JULIUS: So in the general process one of the pieces that was missing from these peer review comments was that many of the plants had not done the -- documented the screening that was done and identification of the pre-initiator. So now we have it in the software, the ability to put in screening criteria and list the surveillance and test procedures and indicate which screening criteria were applied. That's all this shows.

CHAIRMAN APOSTOLAKIS: I mean, if I look at the front picture there, what do I learn from that? Take one entry and tell us what it means?

MR. JULIUS: Okay.

CHAIRMAN APOSTOLAKIS: Anyone you want.

MR. JULIUS: All right. So we have a component cooling water system annual test.

CHAIRMAN APOSTOLAKIS: Okay.

MR. JULIUS: This one right here. And then we list different criteria. And we structure this in a hierarchy to promote defensibility. For example, if components are being tested, it's not in the PSA model, that's the easiest and clearest way to screen it.

CHAIRMAN APOSTOLAKIS: And there is a reason why it's not there, right?

MR. JULIUS: That's right.

CHAIRMAN APOSTOLAKIS: Okay.

MR. JULIUS: If it's in the PSA model, it is not relevant to the top event; then that's our second criteria. For example, if it's a containment system that doesn't link into the LERFTOP.

And then the bottom one would be if it's an insignificant contributor to the PRA results. So we don't like to use that one because it's difficult to defend and you could become in different configurations or conditions where you'd have to reprove that. So we --

CHAIRMAN APOSTOLAKIS: Is it possible that it may become significant?

MR. JULIUS: It is. So that's why we say -- we recommend --

CHAIRMAN APOSTOLAKIS: I don't understand this. You say you don't like to use that, yet it's there. Why don't you take it out? Somebody else insist that it should be there?

MR. JULIUS: Some users will use it, yes. And it's our recommendation on what's a way to do the screening and then when to use it, when not use and it's up to the user then to select what they would like to use.

CHAIRMAN APOSTOLAKIS: Look up this number six there, procedure of deficiency. What does that mean?

MR. JULIUS: The bottom set primarily came out of a review of the historical data. That this is, in this case, something that was found in the procedure, either like the work package was written wrong for installing something or the surveillance and test procedure had a deficiency.

CHAIRMAN APOSTOLAKIS: No, wait a minute. Wait a minute. I mean, say it was found. I don't believe that when you do an HRA you're go and check every procedure, whether it's correct or not?

MR. JULIUS: No, no. This is, as I said, the historical screening of licensee event reports. If there's a licensee event report that said that the condition was found and that the root cause of this valve being found out of position or these instrument were miscalibrated wrong was that the procedure didn't account for the type of calibration equipment or that there was --

CHAIRMAN APOSTOLAKIS: These are so-called latent errors, right?

DR. ELAWAR: Correct.

MR. JULIUS: Yes.

CHAIRMAN APOSTOLAKIS: Slipping there.

MR. JULIUS: Yes.

DR. ELAWAR: Correct.

CHAIRMAN APOSTOLAKIS: But the models that are in the Calculator do not deal with latent errors, do they?

MR. JULIUS: They do in both.

DR. ELAWAR: Yes, they do. The pre-initiators. The pre-initiators are latent errors that lay dormant until --

CHAIRMAN APOSTOLAKIS: Well, the pre-initiator and latent are not the same thing. I mean, pre-initiator means during a test they make a mistake. Latent means that it's buried there someplace and it will --

DR. ELAWAR: That's a pre-initiator.

CHAIRMAN APOSTOLAKIS: They are. They are.

MR. JULIUS: That's part of the screening process. We're identifying these pre-initiator errors that become latent and that will effect the PRA results and should be included in the PRA.

CHAIRMAN APOSTOLAKIS: Some of them. Some of them.

DR. ELAWAR: They will not be revealed until suddenly you need them --

CHAIRMAN APOSTOLAKIS: Do you have any idea how often we find procedural deficiencies?

DR. ELAWAR: Well, that's a good question.

CHAIRMAN APOSTOLAKIS: I mean, we're talking about it, but does it make any difference to the numbers.

DR. ELAWAR: I mean, are we giving certain weight to the possibility that there is a procedural deficiency?

MR. JULIUS: I don't think so. No, no, no.

DR. ELAWAR: This is only showing the comprehensiveness. I have never had a case where I'd say yes, we have bad procedures, here before I would take a higher value. That's not how it works.

CHAIRMAN APOSTOLAKIS: You can't defend that. Even if you want to say, it's difficult to do that.

DR. ELAWAR: I know. And nobody's saying. This just shows the comprehensiveness of the guideline we see here.

CHAIRMAN APOSTOLAKIS: Well, I don't understand how something can be comprehensive if it's irrelevant to the model later. I mean --

DR. ELAWAR: I don't know of any --

CHAIRMAN APOSTOLAKIS: It shows that --

DR. ELAWAR: It happened before, that's all it's saying. And if I'm doing a work here --

CHAIRMAN APOSTOLAKIS: But isn't that half of the model here? I mean, Idaho did studies a few years ago, I don't know if you're familiar with it, where they found that a significant number of errors could be classified or I don't know whether the error or itself or its cause, could be classified as latent. And I don't think we're doing much about it, actually. But maybe that's certifying one that will come later. I mean, I'm not asking you to solve the problems that we have now.

MEMBER BONACA: Well, I'm trying to understand out here this --

CHAIRMAN APOSTOLAKIS: I don't think it's used, Mario.

MEMBER BONACA: When you got to this

component cooling you're trying to find what's the likelihood that in performing that inspection, okay, the operator, the equipment operator will leave something behind. Okay. That's the reason you attempt here to do. And then that's why I'm confused with the procedure of deficiency.

I mean, I understand if there was a procedural deficiency that may lead him to leave something behind --

MR. JULIUS: No. No.

CHAIRMAN APOSTOLAKIS: Ah, we have a problem. Can you hear him? No. We need a microphone.

MR. JULIUS: So there are two separate pieces here. This is the procedure screening on this screen and the resolution isn't very good. So these are surveillance tests.

MEMBER BONACA: Okay.

MR. JULIUS: And normally these bottom three or four wouldn't apply.

MEMBER BONACA: Okay.

MR. JULIUS: Then our good practice is not only to review the procedures, but it's also to look at historical data. Because historical events happened that in spite of the best intended procedures and the best training, things happen. So we look at licensee event reports. And we find in cases where something has happened, an event, a utility will say that this was attributed to a procedure but we fixed the procedure. So that event should be screened. And that's one approach that's been taken.

The supplementary approach that we've advised is that well maybe that should be taken and you should consider for screening, but you should also consider for incorporation of the model. Because if there's something related to that particular component or that environment, or the test equipment they're using that is related to this procedural deficiency, you might generate future ones in that area.

CHAIRMAN APOSTOLAKIS: All right.

MR. JULIUS: So this was our generalized criteria here on the left. And then sometimes they apply to the procedures, sometimes they apply to historical events.

MEMBER BONACA: Okay.

CHAIRMAN APOSTOLAKIS: Okay. Next.

MR. JULIUS: All right. The next few slides are indicating the basis event data, generalized event data that are collected in various screens in the Calculator. The bottom left summary here says it all. This is qualitative data that is common regardless of which method you're choosing. And so we collect it and then combine it differently depending on the method you're using.

So we go basic event data, such as the event name and the description, what procedures are being used, how often they're done, what's the period of testing.

And I'm going the wrong way again.

The performance shaping factors, these primarily come from ASEP. This is the equipment configuration, the I&C layout, the quality of written procedures and the quality of administrative controls.

CHAIRMAN APOSTOLAKIS: Would you walk us through a branch there of the tree?

MR. JULIUS: Sure. So if the highlighted branch there is if we have a good equipment configuration and the I&C layout is good, the quality of written procedures is good and administrative controls is good, that the basic human error probability is 3(e)-2.

CHAIRMAN APOSTOLAKIS: No. How many utility analysts do you expect to say that these are no good? Has anybody ever from any utility say no my quality of my procedures is poor?

I mean, what is this? This is just --

DR. ELAWAR: The configuration is poor. I could have some cases where I could --

CHAIRMAN APOSTOLAKIS: You could, has anybody ever done it?

MR. JULIUS: Yes, sir.

CHAIRMAN APOSTOLAKIS: They've sat and done it?

MR. JULIUS: Well, the case where they do go back to these trees, and typically not in the look ahead. In the retrospective when we get into the significance determination factor --

CHAIRMAN APOSTOLAKIS: Oh, retrospective. But prospective, but I doubt that anyone will say --

MR. JULIUS: That's right.

CHAIRMAN APOSTOLAKIS: -- that I have something poor. So I don't know how useful that tree is for prospective analysis. For retrospective, yes, sure.

MR. JULIUS: We have seen similar trees with similar questions for the post-initiators. And when we have cases when we've gone through and done this type of analysis and we've gotten the feedback from the people performing the procedures or the operators that says, yes, we've got this -- this procedure in general is written well but for the scenario you described, we have these kinds of questions. When we find those things, we use that as a feedback mechanism to make the written procedures better.

CHAIRMAN APOSTOLAKIS: Of course if you find anything, presumably you find it.

MR. JULIUS: That's right.

CHAIRMAN APOSTOLAKIS: So you always end up with good, which is not bad.

DR. RAHN: But it makes people explicitly think about that you have to have good procedures.

CHAIRMAN APOSTOLAKIS: I understand there is a contribution there. But it seems to me that trees like that are really not helpful in prospective analysis. Because I don't expect anyone to say, hey, my plant has bad procedures so I will put a factor there to increase the failure rate. Come on now, let's be realistic.

Let's move on to the next slide with this happy note.

MR. JULIUS: Okay. Then ASEP is a development from THERP and follows a similar, a tasked based or identification of the critical steps and the potential for recovery. So in the Calculator we have one screen for the documentation of the critical steps. For example, failure to open -- reopen a manual isolation valve. Then we look at the factors that are affecting recovery. Is there a compelling status indication, an effective post-maintenance or calibration tests, independent verification or a status check daily or --

CHAIRMAN APOSTOLAKIS: Jeff, I'm looking at the last column there. It says basic HEP three ten to the minus 2, is that what it says?

MR. JULIUS: Yes.

CHAIRMAN APOSTOLAKIS: And then recovery it says one? What does that mean? That if you follow this branch --

MR. JULIUS: That this branch right now has no recovery applied.

CHAIRMAN APOSTOLAKIS: Are these numbers referring to one branch, the red branch? Probably because you give media, mean --

MR. JULIUS: Yes.

CHAIRMAN APOSTOLAKIS: So recovery of one means what? That it will not be recovered. It's a failure probability, right?

MR. JULIUS: That's right.

CHAIRMAN APOSTOLAKIS: There's no recovery? And what's the difference between basic HEP and mean value of the HEP?

MR. JULIUS: On several of the NUREGs the HEPs were listed as medians and we did the median to mean conversions. Some utilities have consistently used medians and some have adopted converting the values to means.

CHAIRMAN APOSTOLAKIS: So this particular one uses the basic as median?

MR. JULIUS: And we show both the median and the mean there.

CHAIRMAN APOSTOLAKIS: No. But this one uses the basic the HEP as the median, right? Three ten to the minus 2, three ten to the minus 2?

MR. JULIUS: Yes.

CHAIRMAN APOSTOLAKIS: So basic refers to some document 1278, or something?

MR. JULIUS: The 4550.

CHAIRMAN APOSTOLAKIS: Okay.

MR. JULIUS: The ASEP dependency factors are the actions close in time and the same visual frame of reference, same general area. Is there writing down required. So this is the probability of A and B. They are in close in time, yes. And in the same visual frame of reference. Yes. Then the level of dependence is complete.

CHAIRMAN APOSTOLAKIS: Are you in the quantification then, how do you handle a level of dependence? Are you going to talk about it?

MR. JULIUS: This is where we talk about the quantification for the level of dependence in the pre-initiators. So this would be a --

CHAIRMAN APOSTOLAKIS: Do we have another slide later or should we talk about it now?

MR. JULIUS: We have another slide later for the post-initiators between our reactions.

CHAIRMAN APOSTOLAKIS: How do you handle these in the pre-initiator? I mean, what do you do to the probabilities?

MR. JULIUS: Oh, we take A and B; A as the base HEP and B as the recovery probability. We would adjust the recovery probability B to be a conditional probability based whether it's qualitatively low, medium, high; they map to using NUREG-1278 to be 1 plus 19 N over 20 for the low dependency.

CHAIRMAN APOSTOLAKIS: Oh, you are using those?

MR. JULIUS: Yes.

CHAIRMAN APOSTOLAKIS: You notice the long silence?

MR. JULIUS: Yes.

CHAIRMAN APOSTOLAKIS: Okay.

MR. JULIUS: THERP is --

CHAIRMAN APOSTOLAKIS: I tell you, those if you think about it, they always give you one or two numbers. I mean, the formula is misleading. Because there is --

MR. JULIUS: That's right. It's a .5 or .05 of .16.

CHAIRMAN APOSTOLAKIS: Yes.

MR. JULIUS: Or the base or one, yes. It's essentially five values. I have thought about it.

So the pre-initiator or the third method, this is where again we're talking a look at the critical steps. So this slide just shows the step number and instruction. And it shows the errors of omission, a commission table that you would select from THERP, but it's a similar type of approach.

When you use the software it shows the tables here on the left, the THERP tables are linked in. And then when you select the item from the table, you can easily see and go through the checklist. Is this an analog meter with easily seen limit marks or a digital meter?

The THERP approach does allow for multiple errors of commission. For example, the misreading or failing to hold the switch over as well as selecting their own switch.

This is our graphical display of the THERP critical actions and the recoveries. So we list all the steps that are done and then we typically apply one of the steps such as open a valve and then later on check that the valve is open. We showed in this case that it's assessed with a low dependence, again, using a similar type of approach to the definition of the dependence level.

And then the THERP summary, what you see here is that the critical steps, the recovery steps, what are the actions and the level of dependence, what the total error is and then what the different contributions. So, for example, on these event the 5.90 minus 4, the biggest problem is coming through the reconnecting the pump there and 7.10.5, 2.60 minus 4 out of the 5.90 minus 4 is coming from that steps. So it allows you then to look back at what is driving the results as well as the total.

CHAIRMAN APOSTOLAKIS: So what you have done is you have developed the software tool that helps a user of the THERP method for pre-initiator errors, help the user to use the 1278, essentially, NUREG-1278, right?

MR. JULIUS: Correct.

CHAIRMAN APOSTOLAKIS: This is very useful.

Have you changed in a significant way any of the numbers in that document or have you simply computerized it?

MR. JULIUS: For the pre-initiators we've simply computerized.

CHAIRMAN APOSTOLAKIS: Okay. Good.

DR. ELAWAR: Changed from medians to means.

CHAIRMAN APOSTOLAKIS: You have changed --

DR. ELAWAR: We are using means --

CHAIRMAN APOSTOLAKIS: But I think Swain made it clear that his best estimates were median.

DR. ELAWAR: Well, the industry is using mean values all throughout.

CHAIRMAN APOSTOLAKIS: Well, you can use mean values if you did IVAN.

DR. ELAWAR: Yes, we did IVAN in the Calculator and we used that.

CHAIRMAN APOSTOLAKIS: What Swain and Gutman say, they give you a best estimate and two bounds, right?

DR. ELAWAR: Correct.

CHAIRMAN APOSTOLAKIS: And the three of them are consistent with the log normal distribution. They are consistent. So the middle value is the median and the others have the fifth and the 95th. So now you're saying, no, the median -- what he says is the median we will use as a mean?

DR. ELAWAR: That's what we are saying.

CHAIRMAN APOSTOLAKIS: Well, that's not right. I mean, if a guy says best estimate is median, I mean you should respect that. If you want to use means, divide it. You can divide it easily.

MR. JULIUS: We have two general camps within the EPRI users group. One is that, yes, it's listed as a median and it says the error factor and here's the way to mathematically convert it to means. And in general, the ASME standard promotes means, so those conversions have been done. And the other that it said that our level of knowledge between the median and whether it's a median or a means is the centralized best estimate value and we use the medians directly.

CHAIRMAN APOSTOLAKIS: The mean. Yes. On the other hand there is strong evidence that the expert judgments, even if the expect claims that he's giving you his mean value, he's really giving you the 50/50 value because our brain doesn't work that way.

The mean value as well as the variance are mathematical occupiers. Our brain doesn't integrate and get a mean value. Usually we work with -- I'm surprise that you guys are doing this. But other than that, I think it's a good thing to do.

DR. ELAWAR: Yes, that's a consensus. And I agree with you, it can go either way. But the -- was different to do those as medians and convert to means and use that.

CHAIRMAN APOSTOLAKIS: You know, in the original draft of 1278 --

MR. JULIUS: Yes.

CHAIRMAN APOSTOLAKIS: -- the bounds and the best estimate were not consistent with the log normal distribution, and there was a major comment and Swain changed it. So it's not something that he did on the side. I mean, it was something that he thought about. Swain and Gutman thought about it and they're telling you these are, you know, the advice of a long and normal distribution. I mean, I don't know how you can take liberties with that and say no, no, no. You guys who wrote the 1,000 page report don't know what you're talking about. You are giving us something else.

Anyways, shall we go on?

MR. JULIUS: Let's. Switching gears here now to the post-initiator model. When we get to the end, I'll reanswer your question on what we've changed with respect to the values and the base reports.

The approach is very similar here. You can see it on the far left of the screen. These are the basic steps as we step through the different aspects.

We start with the basic event data. What's the label for? It's a description. We fill in the different cues and indications. And we've left sufficient field and room here for the primary cues, secondary cues as well as additional indications.

The procedures, list the procedure for both the cognitive and execution and the types of training. Is it trained in the classroom, trained in the simulator and at what frequency or is there a job performance measure that's associated with this action?

The scenario description, you see from the screen, we've left it as one big blank text box. So in general from a software point of view it's a free formatting field that you could put whatever data you want in there. From the user group's perspective we have looked at different human reliability analyses such as Palo Verda's and several other plants and have combined a Best Practices. We suggest when you're doing the evaluation of the scenario, that you consider the initial conditions, the initiating event, what's the accident sequence, the preceding functional successes and failures, what's the operator errors that are part of this sequence, what's the success criteria for this action, what's the consequences of failure and consequences of success? So we lay out a practical comprehensible approach to defining this area. And it allows also for documenting then the inputs from the operator interviews or from simulator data.

Here's the time window that I was describing with the overall time on the top. That's the system time window available for action before the universal damage state. And then we breakdown the lead up for the action; that there's some time delay, then a cue occurs. And after the cue there's this cognitive processing and manipulation. The manipulation time includes both the time to manipulate the valves as well as any time to go out if it's a local action, to get to the area of transport time, for example.

And then from this time window at the end then we see the time that's available for recovery. So if we subtract off all these time that's used up at the beginning and then we also list on there the SPAR-H, the available time both cognition and execution. One of those is a difference and one of them is actually a ratio. So the difference between the system time window and the time that's been used up, for example here on the slide here, that's 82 minutes is remaining for recovery. And then a ratio method, this 82 minutes and there's about 8 minutes needed for the manipulation. So you could do the manipulation 11 times.

CHAIRMAN APOSTOLAKIS: I don't follow. The first time 82.3 it says there?

MR. JULIUS: Yes.

CHAIRMAN APOSTOLAKIS: That's minutes and it comes from thermo-hydraulics?

MR. JULIUS: That's the -- no. The system time window, it typically comes from a thermal hydraulics. And what we've chopped off here is the ability to link to the thermal hydraulics.

CHAIRMAN APOSTOLAKIS: But it's 120? What is it?

MR. JULIUS: That's 120 minutes for this

example.

CHAIRMAN APOSTOLAKIS: So why do you say then the time available for recovery is 82 minutes?

MR. JULIUS: This is for recovery of the first action. Because it takes in this case, there's 30 minutes of delay and 8 minutes to do the action initially.

CHAIRMAN APOSTOLAKIS: Yes.

MR. JULIUS: So there's 38 minutes just getting to it and through it the first time.

CHAIRMAN APOSTOLAKIS: And then you realize that something is wrong.

MR. JULIUS: And then this is how much time is now available after that for recovery of that first failure.

CHAIRMAN APOSTOLAKIS: Assuming it was not caught earlier.

MR. JULIUS: Assuming it was not caught earlier. And some of that could be not caught because you were doing other things or because you made the mistake, even the cognition or the execution.

CHAIRMAN APOSTOLAKIS: Yes. Okay.

MR. JULIUS: And that level is used later. I'll show that.

CHAIRMAN APOSTOLAKIS: So you are using

some stuff from SPAR-H?

MR. JULIUS: Well, we use that as a feed to the SPAR-H. Again, we're collecting this qualitative data and then we're using it in the different types of methods.

When we very first put it down, the time window documentation and definition was different between HCR and caused-based decision tree and SPAR. And we said no, we need the analysts to have a simple common picture for the timing.

So if you were using this for SPAR, then that was for the timing data.

CHAIRMAN APOSTOLAKIS: I thought SPAR-H was not one of the models?

DR. ELAWAR: This doesn't mean that analyst use. This is just for a reference in case he wants to compare it with SPAR-H. That doesn't mean it's being used in the actual EPRI analysis. It's just he put it here in case I want to compare later on, I will have things available to me. But the bottom line --

MR. JULIUS: Yes. There's no possibilities there. One is that, again, an analysis of an event such as the significance determination, a local SRA or somebody might call up and say we did a SPAR analysis on this event and we come with a factor of recovery of X and the utility guy says well I came up with Y. And when we're looking for differences, this will allow them to talk in common terms of what kinds of time are you seeing available for recovering using SPAR.

I've also had one of the vendors was talking about using SPAR as a look ahead for some of the initial quantification of their human interactions.

And this part might be new to some of you, in that the cause-based decision tree method, this is an EPRI proprietary method in that it was developed through EPRI research funds.

What we see here is that there are eight different decision trees, four of them having to do with the man/machine interface and four of them having to do with the way the procedures interact. And it questions things like availability of information, failure of attention, misread or miscommunicate data, skipping a step in the procedure or misinterpreting the instruction or having a tough decision logic. So we picked one those of trees, the availability of indications and shown graphically how we step through the tree and then have fields to allow for the documentation of the notes or assumptions when you're doing that event.

CHAIRMAN APOSTOLAKIS: Has any utility submitted a PRA that did the human reliability analysis this way to the NRC?

DR. ELAWAR: Single items, yes. But not a whole report.

CHAIRMAN APOSTOLAKIS: Single items means what?

DR. ELAWAR: Because we have an SBP case and we need to redo an HRA, we do it by the HRA Calculator and we'll submit that information.

CHAIRMAN APOSTOLAKIS: And what does the NRC staff say?

DR. ELAWAR: As far as know, use the Calculator has never been rejected in terms of adequacy of HRAs. I have one example for example for you. I have a Calculator one HRA value and compared with what the NRC have done in SPAR-H. Things that I say no I don't take credit for this, because there is no procedure. In SPAR-H they were taking credit for it and I'm disagreeing with it. I'm saying that sometimes that we are more conservative than what SPAR-H allow.

CHAIRMAN APOSTOLAKIS: Well, the issue really here is when you say EPRI proprietary, what do you mean. Has the NRC staff reviewed it? We are hunting proprietary information all the time. Has this been reviewed by the staff?

DR. ELAWAR: It was offered for review, am I right, some three years ago.

DR. RAHN: It's available to staff, whether or not they have reviewed it I don't know.

CHAIRMAN APOSTOLAKIS: So the staff has access to it? Okay.

MR. JULIUS: I have received comments both from staff or supporters of staff or from people around the world that haven't seen or are not familiar with this approach because of the --

DR. LOIS: This is Elrasmia Lois.

We did. We reviewed CBDTM and it's going to be discussed in the next presentation.

CHAIRMAN APOSTOLAKIS: Okay. Good.

Boy, I like your arrows there. I mean, they're so impressive.

MR. JULIUS: It's part of the human factors for the slide.

CHAIRMAN APOSTOLAKIS: Yes, I know.

MR. JULIUS: So there's a lot of data n this slide, and I was trying to think of a way to easily convey the general meaning here.

DR. RAHN: It's also coordinated with the weather. See, if it wasn't a snow day today, you'd have blue.

CHAIRMAN APOSTOLAKIS: Something else?

DR. RAHN: Otherwise it would be yellow or whatever.

CHAIRMAN APOSTOLAKIS: Okay. All right.

MR. JULIUS: But what I intend to show you--

CHAIRMAN APOSTOLAKIS: Do you what snow is, Frank? In California, do you know what snow is?

DR. RAHN: Yes, I used to know but I've kind of forgotten.

CHAIRMAN APOSTOLAKIS: Something that comes from the sky.

MR. JULIUS: So this isn't something that's coming from the sky. So this is human reliability. And I start out with --

CHAIRMAN APOSTOLAKIS: Does human reliability come from the sky, Jeff? Is that what you're --

MR. JULIUS: Some perceptions are, yes, sir, is that it does.

CHAIRMAN APOSTOLAKIS: Divine perceptions.

MR. JULIUS: So we have on the left side

here the cause-based decision tree that produced the contributor. In this example we have PCB, which was the failure of attention and skipping a step and having trouble interpreting the logic. So that's PCB/PCE and PCG. And then we look at the different recovery factors available; self review, STA review, shift change and ERF.

This is one of the places where the Calculator does some suggestions that help improve on what you would find if you were just picking up the report. If you were picking up the report, you'd see this matrix up here, these different factors available for recovery and you could select, for example, multiple factors. You could theoretically on this PCE you could pick extra crews, self-review, shift change or ERF review. We know from the timing data that was in put previously, you can see in the upper right hand slide that the time window was 120 minutes and there was 82 minutes available for recovery. Because there was only 120 minutes from time zero, we don't credit or allow with the software credit for shift change or the ERF review depending on the timing. If it's too short. So we take away those possibilities.

And we also suggest -- we limit the operator to pick the best recovery mechanism. Is it self-review or is it extra crew. Because there have been a tendency in former HRAs to pick as many as you could. Okay. I've got three that's available, I should do three. And if you appoint one three times, then all of a sudden you have factor of 1,000 applied and things disappear.

Also the timing in this case we have 82 minutes available fore recovery so we have plenty of time before recovery. We have a little diagram that shows if the timing gets restricted that you should say that the recovery factor is limited to a high dependency, for example, or a moderate dependence. And that's what I've shown here on the arrow two going to the dependency factor column. That if you had a case where you had maybe 20 minutes available for recovery, that a moderate dependence should be applied. And instead of using a 1.1 or 5(e)-2 then you would in this case a .16.

And these are summed across and down to give the cognitive portion for the cause-based decision tree.

CHAIRMAN APOSTOLAKIS: These are all point estimates, right?

MR. JULIUS: Yes.

CHAIRMAN APOSTOLAKIS: There is

uncertainty later, uncertainty evaluation?

MR. JULIUS: That's right.

CHAIRMAN APOSTOLAKIS: Okay.

MR. JULIUS: And for the execution portion, that's the cognitive and there's performance shaping factors and stress. The stress was one that was questioned earlier.

The upper left screen is the general qualitative performance shaping factors; the environment, the lighting, humidity, heat, radiation, atmosphere. Are there any special tools, parts or clothing required. What's the accessibility of the equipment.

Then you see for the stress is the plant response as expected, yes or no. Is the workload high or low. And then a separate button for the performance shaping factors being optimal or negative. And this is a case that I know present John Forester hasn't seen before where the previous answer is here. For example, if you're in emergency lighting or if you're at a hot humid environment or a smokey atmosphere, the inputs on that previous screen will then indicate that you've got negative performance shaping factors which would tend to drive the stress level up. This was a recent addition or improvement that we've made.

Okay. Then we jump over. And this slide is meant to show the cognitive human error probability that comes using the human cognitive reliability, operator reliability experiments.

In this approach the timing data implicitly includes the performance shaping factors. And that typically comes from operator interviews. And it's important then and we stress that when you're getting this timing data from the operators, that you need to discuss the progression of the whole scenario. If you call up and ask an operator "Hey, how long does it take to do this?" He can do anything in five to ten minutes and there's always success. So it's like okay, let's start from the beginning. What are you seeing here and how long it does it take. When you're going through these different steps, what steps are done parallel, what steps are done in series and what's the full progression. Because there's a tendency to forget some of the time delays or the distractions that involve getting to the point where you've got that five to ten minutes.

The HCR/ORE approach then also has the other primary variable, the evaluation of sigma, which is the variation between the crews. We have the ability for people to develop their plant specific data for the sigma. And we provide a simple decision tree approach for the variation of the crews. You can also get it from the EPRI experiments that were done previously.

CHAIRMAN APOSTOLAKIS: But sigma is not representing only crew to crew variability, right? I mean, I thought it was uncertainty about the time. It's the sigma of the level of the distribution of the time, right?

MR. JULIUS: That's right. But it's also meant to collect the variations of the crew.

CHAIRMAN APOSTOLAKIS: It may include the crew to crew variability.

MR. JULIUS: Yes.

I skipped over showing the third for the execution because it's the same process that was used for the pre-initiators; there's the critical steps recoveries that are applied, look up tables that are included in the software.

Then I've gone back to the main scream there for the basic event data. And what we show is that the contribution from the cognitive with and without recovery, the contribution from the execution portion with and without recovery and the total human error probability. So you can drill back from the total human error probability is that primarily cognitive or execution driven and what's the different factors.

What you don't see here is that the tool then also provides ability to do this consistency check so we can print out. Because all the information is in a database; the list of the human error probabilities, the basic event ID and some of these different factors is it high stress, what's the timing and so you can line them up and then qualitatively say well that makes sense. This one has a higher human error prob ability because there's not much time available, it is a higher stress. And it's just a cross check that can be done.

One new feature looking ahead for 2006, because it is that time of year, is that one of the utilities says they have plant specific data for their cause-based decision trees which was encouraged in the EPRI report. And they want the ability to put their own data in for the cause-based decision tree. So we're looking at adding that for 2006.

The one thing we've done, the feature with having this in a software approach is that now that for this operator action and using this method, this qualitative data can then be opened up when you open up -- if you say for example you start with the cause-based decision tree approach, you open up the human cognitive reliability, all the qualitative data and the timing data is there. You would add any new factors such as the sigma and you could see what the results would be then using a different method.

You've asked about the uncertainty. Well, we have the error factor is primarily derived from the total human error probability using that simple table from Swain basically says if it's a low human error probability we give it a bigger uncertainty factor and if it's a larger human error probability, it's smaller. But the approach we've taken is that a lot of these factors can be driven by some of the assumptions, either the method that was chosen or the selection of the stress, for example, or maybe some variations in the timing values. So with this tool you can then save this case and evaluate several sensitivity cases to get a better feel for what is, for example on lower bound or upper bound, on the human error probability.

CHAIRMAN APOSTOLAKIS: The error factor essentially is assigned independently of what you did. I mean, you said you used Swain's --

MR. JULIUS: That's right. We don't take the Monte Carlo or roll the different error factors for the different things up into the total. We just say look at the total and then assign the error factor based on that.

CHAIRMAN APOSTOLAKIS: Okay.

DR. ELAWAR: And if I may add a comment here? That is a little bit more than that. I usually, and I know my peers also do, the sub tasks in each qualification from say THERP have error factor with them. When I look at them at the bottom of my error factor I compare with sub task and make sure that there is reasonableness in it, without necessarily applying Monte Carlo techniques for it.

CHAIRMAN APOSTOLAKIS: But if you have dependencies, for example, and you use the formulas that are handle says, it seems to me a major source of uncertainty is the validity of the formula itself. So you really have to at the end judge what you have included in your calculation and what's the uncertainty.

DR. ELAWAR: That's a valid comment.

CHAIRMAN APOSTOLAKIS: Which contradicts your earlier statement for the uncertainties here are the same as those for the hardware.

DR. ELAWAR: Not the same. There are a variety of uncertainties for the sub tasks, and I want to make sure I'm not totally out of range with the sub tasks.

CHAIRMAN APOSTOLAKIS: Okay.

MR. JULIUS: And before we move on, the next section of the presentation there's a short description on the dependency between human interactions.

One of the differences between this approach and, for example, SPAR or ATHEANA is this lays out, for example in the cause-based decision tree, it gives a standardized checklist of here's the cognitive, eight ways or potential failure modes. It's hardwired and set that those are eight and you see the different ways those can fail. ATHEANA takes a step back and says well are there other questions that should be asked. This is probably more valid again in the retrospective review. I think in the prospective look or application of ATHEANA there'd be a tendency to fall on well when we're looking ahead there are a standardized set of here are the typical questions it asks and it's more difficult to anticipate. For a prospective should there be something else that is asked.

And then on a comparison with SPAR, by going with the caused-base approach and looking at the tasks and the failure modes and the recovery, we've taken it another level of detail down below what SPAR typically asks. SPAR typically in general is there adequate time, expansive time, what's the procedures in general. And you don't see the link. You know, is the fact that the procedures are trained on once every five years or that the procedures have a wording problem. That comes through clearer here in the Calculator and the approach that we've taken.

We do have the worksheets from SPAR-H for both the cognitive and action. And you can see --

CHAIRMAN APOSTOLAKIS: You take their numbers, you take their worksheets but you still maintain you're not using SPAR-H?

DR. ELAWAR: Correct.

CHAIRMAN APOSTOLAKIS: Okay.

DR. ELAWAR: And there's no law against it.

CHAIRMAN APOSTOLAKIS: I know there's no law. But there ought to be one.

MR. JULIUS: I would say --

DR. ELAWAR: I knew the fact. But as far as know are not using SPAR-H for their bottom line reporting.

CHAIRMAN APOSTOLAKIS: They're using the Calculator?

MR. JULIUS: It's not used in the prospective looking at here's the evaluation of our HRA update. It is being used in the evaluation of individual events involved with the significance determination process.

CHAIRMAN APOSTOLAKIS: Just to be prepared for that.

DR. ELAWAR: It make sense.

MEMBER BONACA: How do the evaluation with HRA compare to the one with your two?

MR. JULIUS: How does the SPAR evaluations compare?

MEMBER BONACA: Yes.

MR. JULIUS: We haven't conducted that exercise yet.

MEMBER BONACA: Okay.

MR. JULIUS: I know in the SPAR-H they go through and they document their comparison using THERP and several other standardized approaches. They've done a consistency check that way. But our members are just starting to ask for that type of look ahead.

MEMBER BONACA: But wouldn't it be

important or interesting? I mean, at some point for the utilities if they have been evaluated on the basis of SPAR-H evaluations, you would want to know how well you're agreeing with estimations.

MR. JULIUS: Right. And that SPAR-H report was published August of 2005. So that was --

DR. RAHN: It takes a while. There's big quality assurance steps that we have to go through before we are ready. But, yes, we are going in that direction and that is important.

CHAIRMAN APOSTOLAKIS: Okay. Can you speed it up a little bit, Jeff?

MR. JULIUS: Yes.

CHAIRMAN APOSTOLAKIS: We talked about it, didn't we?

MR. JULIUS: The next few slides are the dependencies between human interactions. The development of the events you've seen so far were the dependencies within human interaction. So the generalized approach as searched with the human failure, identification and qualitative definition, it's addressed during operator interviews. And then what's of most interest lately, is the double check with the quantification results. So we're looking at the cutsets or the sequences and then evaluating the level of dependence and readjusting the logic model accordingly.

So the recent feature of the Calculator is the ability to import cutsets. And then what you see here in the upper left is cutset number 1 is a combination of hardware and initiator and human error reactions. And you can see in this example there are two human interactions that are in the model. And the parameters here are the individual probability for each and then the timing factors that are involved. So the system time window, the time delays in the manipulation. And this way you can see whether they're occurring close in time or not.

If you want to drill back out and see what types of initiators are involved, that's what the bottom right screen is showing, that this pump that's incut set number 1 is showing up in the general transient as well as loss of instrument error cutsets but it also has these -- for the general transients it has these hardware contributions. And for the loss of instrument error, it has these other hardware contributions. So we're trying to make it easier to identify those combinations and the scenarios that they're involved in.

We have interfaces, more ability to

combine databases and we export then the results directly into NUPRO or CAPTAFILE for use in the quantification. When that export process is done if a human error probability is quantified to be below a user defined value of say ten to the minus 4 or ten to the minus 5, then it's imported as ten to minus 4 or ten to minus 5, it doesn't import as ten to the minus 12 or 13.

Each event then is documented in a written report for that individual human failure event. Again, the qualitative factors as well as the quantification.

And that's the technical description for the HRA Calculator.

CHAIRMAN APOSTOLAKIS: Thank you.

Who is doing this?

DR. ELAWAR: You want to do it? I'll do it?

CHAIRMAN APOSTOLAKIS: All three of you guys. All three stand up.

I mean, we have extra chairs, don't we? Yes. All three of you can sit up front there.

DR. RAHN: Just in conclusion, Mr. Chairman. Again, thank you for inviting us here. We did want to make a few observations.

First of all, industry --

CHAIRMAN APOSTOLAKIS: Jeff, pull up a chair.

Okay.

DR. RAHN: Industry has recognized a number of years ago that there were inconsistencies in approach and whatnot. And the purpose of the EPRI program is to solve those, and we've been working five years to improve the ability of users of the utilities to do HRA. We believe most of the prior deficiencies have been corrected, but again our mission was to develop a tool that was widely accepted, uniformly applied and a transparency so that we understood the strengths and the weaknesses of what we were doing.

We believe that the Calculator approach satisfies the standard, the ASME standard. And we work also to ensure that it meets the NRC Good Practices for implementing HRA.

Right now the industry believes it meets its needs for its safety analysis and for its regulatory needs. And that, of course, was the important thing that we needed accomplish.

We are moving to go beyond PRA level 1, which is internal events, shutdown others are the types of things we're working on. And we try to monitor the research done by others, including the NRC and our international partners.

We are adapting a fairly conservative approach in terms of implementing new models. First of all, we need to have the transparency, the traceability, defensibility, the useability. That is, we recognize that we have a need to train our users in what we're doing. And unless a new procedure is well documented as, if you will, gone through the test of time, is well understood we're a little bit slow to implement it for those reasons.

CHAIRMAN APOSTOLAKIS:

I believe that the issue of consistency is very important. And I think having a tool like this is certainly a good step forward. But I still think, though, that you would make a better case if you run some sort of an exercise where you had two, three, four different groups; utility people, you know the way you want the group to be. Give them a sequence or an event, preferably a sequence, and ask them to use the Calculator anyway they want and see what you get. You will get a lot of insights from that.

DR. ELAWAR: (Off microphone).

CHAIRMAN APOSTOLAKIS: You have to speak to the microphone.

DR. ELAWAR: Most likely we'll do that.

I'll introduce this issue to our group meeting --

CHAIRMAN APOSTOLAKIS: Very good. And before you do that, please read that paper from the PSA conference.

As we all know here, of course, we will never have an experimental validation of these models in the sense that, you know, natural laws are validated. We will have to rely on people's judgments and in direct evidence, you know, simulators and all that. So at least trying to achieve some consistency and eliminate a lot of the -- well another insight from the European Union exercises was because they didn't do only the HRA, they did fault trees. I mean, at that time they were new, of course.

A major insight was, which is not surprising to us now, was that the major reason for the discrepancies was that different people used different definitions, different boundaries. Different, not necessarily assumptions, but it was a matter of interpreting what they were supposed to do. And I think that having a tool like this will probably go a long way towards eliminating a lot of those, but I recognize for you guys to demonstrate that and say, yes, we did this, this is what we found as a result of that we're happy or we're changing it a little bit. I really think it would be a great idea to do that.

DR. ELAWAR: Yes. That's a good comment.

DR. RAHN: And that's a good comment, Mr. Chairman.

I might add that in addition to what you've just said, it's the fact that we are training people to the common --

CHAIRMAN APOSTOLAKIS: Yes, that's a value. Yes, you have it on your next slide there.

DR. RAHN: So as we have mentioned, we are training a dedicated core of utility analysts in these methods. We support university research. We have a training package which in addition to our normal training exercises which, like I mentioned, occur about three times a year for, if you will, self-training. That's essentially a five day training course which we have developed in conjunction with our risk and liability usage groups where people can essentially go off and self-train. And that's to the INPO standards.

We have comprehensive sort of guidelines which will compliment the ASME PRA standards. We will automatically link to commonly used PRA tools in the industry. And, of course, we are always anxious to work cooperatively with NRC. We have since we started and always invited NRC personnel to participate in our meetings, and happy to share with the staff any of our research results, etcetera. And we look forward to extending this in the future.

I think that's --

CHAIRMAN APOSTOLAKIS: One last comment for me.

DR. RAHN: Sure.

CHAIRMAN APOSTOLAKIS: I appreciated the discussion we had earlier regarding the models and so on, and Frank points out that you wanted to include models that people have used. But I will repeat that my view is that at some point we have to start saying or advising the user, look, this model is based on very questionable assumptions, period. Don't use it, period.

Now the NUREG draft report that you have not seen doesn't go that far. But at least it's a very good first step when it evaluates things --

DR. ELAWAR: There were some peer review comment in that direction where questioning the methods used.

CHAIRMAN APOSTOLAKIS: Yes, but what's the result of that? Yes, I know that people are questioning. But --

DR. ELAWAR: But since those --

CHAIRMAN APOSTOLAKIS: It's very difficult to tell somebody whom you've know for years that his model is no good. It's very hard. I appreciate that, although people do that to me all the time. But we have to reach a point where we just stop saying, you know, oh here's a bunch of model, you pick, you know.

Any comments from my colleagues? Mario?

MEMBER BONACA: No. I think that I'm impressed with the level of detail, and most of all with these activities that are pulling together the users and providing this kind of training. Because ultimately it's the only way again to achieve some consistency and have, you know, a way of comparing apples and apples between different plans. And particularly from a perspective of the NRC that is working with SPAR as a code to evaluate individual plans and then to quantify in a way that you can compare plans. This provides another help in that direction.

CHAIRMAN APOSTOLAKIS: Tom?

MEMBER KRESS: Well, I think it looks like a good framework to provide this consistency.

I agree with you, George, that an exercise to demonstrate that you get rid of this user inconsistency would be well worthwhile. I think I need to see the database that backs up the actual models. You know, I think it incorporates all the performance shaping factors in a good way, it looks way. And it gives you the options on how to use them. So I'm encouraged by what I see.

CHAIRMAN APOSTOLAKIS: Good.

MEMBER KRESS: But I have to look at the-- you know, you get a number out of and I have to see what the number is based on yet.

CHAIRMAN APOSTOLAKIS: Okay. That's it.

MEMBER KRESS: Yes.

CHAIRMAN APOSTOLAKIS: Well, gentlemen, thank you very much. I really appreciate your coming all the way here to enlighten us. And I certainly was enlightened. I appreciate that. Thank you very much.

DR. RAHN: Well, thank you for your invitation. And we will take your suggestions to heart.

CHAIRMAN APOSTOLAKIS: We'll recess until 10:50.

(Whereupon, at 10:30 a.m. a recess until 10:50 a.m.)

CHAIRMAN APOSTOLAKIS: We're back in session.

The next presentation is by the NRC staff. It will be another view of the human reliability analysis program and we have Mr. Yerokun, Dr. Lois and Dr. Cooper. Please.

MR. YEROKUN: Thank you, good morning.

I'm Jimy Yerokun, I'm chief of the Human Factors and Human Reliability Section in the Office of Research. With me and from my group, my section, Dr. Cooper and Dr. Lois.

Also present or will be present shortly from the Office of Research is Mike Cheok one of the branches in my office.

We have also representatives the folks we work with from Sandia National Lab. We have folks from SAIC and we do have people from University of Maryland. So for the rest of today and part of tomorrow, we'll hope to give you a very good overview the HRA activities we have going on.

When the presenters come up, I'm sure they'll introduce themselves at the time when they come for their presentations.

The objective of --

CHAIRMAN APOSTOLAKIS: Would you tell us a little bit about your background. We know the lady's. It's the first time we see you.

MR. YEROKUN: I've been here a couple of times in the past?

CHAIRMAN APOSTOLAKIS: You have?

MR. YEROKUN: I've been in front of you two times in the past.

I started working for the NRC in 1989. I worked in the original office. I've also been one of the resident inspectors at one of the sites.

I came to headquarters three years ago. I spent a couple of years in the Office of NRR.

Prior to the NRC I worked for the industry. I worked directly for a couple of utilities and I also worked for one of the construction engineering firms.

I've been in the nuclear industry for, say, about 25 years now at various aspects of the industry; construction, startup, operating and with the NRC.

So the objectives are to provide ACRS an update NRC's HRA research program activities. We don't plan to discuss all the program activities, but we definitely have some of those activities selected to give you a little more insights into what we're doing and what the plans we have for those specific activities.

One of our objectives we hope to achieve today also is to obtain some feedback from the ACRS to inform the planning of those activities we plan to discuss today and tomorrow.

We are especially interested in getting some feedback on those activities that are in their beginning stages. That should help us shape the way we move on with those activities.

We also hope today to address some current interests of the ACRS. We're going to add some questions and some of the HRA methods, ATHEANA, SPAR-H. So we hope to be able to address some of those interests.

Just to give a short insight to the goals and objectives of the HRA research program. The goal, we support risk-informed regulatory activities. We have multiple objectives research program for HRA. One of the objectives is to improve existing HRA methods or tools.

One of our objectives in the research program is to provide for technology transfer.

And we also strive to address emerging needs, such as HRA for advanced reactors, HRA capability for a MSS, which this tool is not part of our discussion topics, but those are some of the activities we are engaging with our research efforts in the HRA area.

One of the major focuses of the current HRA research is to support NRC's action plan regarding PRA quality. So we do have ties to the PRA quality program goals. And thus far we have completed the NUREG-1792 which documents NRC's reviews of what the practices are. And you have also the copy of the current draft NUREG that contains some of those existing methods that gives the Good Practices.

And today we plan to present our work so far in this Good Practices and evaluate current methods against Good Practices.

For the briefing overview, we will provide an overview of the HRA program which provides some discussions on some specific HRA program activities and some HRA methods of interest. The HRA Good Practices, the evaluation of HRA methods against the Good Practices. We talk about HERA database and we have colleagues from Halden to present some of our Halden activities. You know, we obviously are very involved with the Halden program.

CHAIRMAN APOSTOLAKIS: By the way, since we have to shorten a lot of amount of time we spend on this, we will be hearing from the Halden people at

4:30 today.

DR. LOIS: Or earlier if we --

CHAIRMAN APOSTOLAKIS: Well, if we finish earlier. So the HERA data and Bayesian methods will be tomorrow morning.

MR. YEROKUN: All right.

CHAIRMAN APOSTOLAKIS: IF it's okay with everyone.

MR. YEROKUN: Okay.

CHAIRMAN APOSTOLAKIS: Since these people are coming from Norway, it's a long way. Okay.

MR. YEROKUN: I appreciate that.

Before I turn it over, I just want to point out that a lot of the activities that will be discussed in the next day or so, we have project schedules to involve the ACRS in those activities at the times that are appropriate. So the intent of today's and tomorrow's briefings would just be overviews, just a broad perspective of efforts in those activities. And we do appreciate the ACRS asking us here to give this big picture view. And it doesn't preclude us from interacting, obviously, in the future or specifically with those activities to get either the approval or the letters of consent from the ACRS as necessary. I just wanted to --

CHAIRMAN APOSTOLAKIS: Yes. I mentioned it earlier to Dr. Lois. We have to schedule meetings with you in the near future. As you know, in February the full Committee will review the comparison with the Best Practices. Maybe you can come back later today or tomorrow and tell us when it would be a convenient time for you to brief the full Committee.

MR. YEROKUN: Sure.

CHAIRMAN APOSTOLAKIS: On other major research efforts you have like SPAR-H and so on.

MR. YEROKUN: Okay.

CHAIRMAN APOSTOLAKIS: So you will get formal letters from the Committee.

MR. YEROKUN: Right. We can do that. That's no problem. All these activities, we have our schedules laid out and at the appropriate times for ACRS interaction, we will come --

CHAIRMAN APOSTOLAKIS: Because I would like the full Committee to also be aware of what you are doing.

MR. YEROKUN: Okay.

CHAIRMAN APOSTOLAKIS: Not just the Subcommittee.

MR. YEROKUN: Okay. Good. Right.

So with that, Dr. Lois.

DR. LOIS: Thank you. I also thank you very much for the opportunity today to discuss our activities and get the early feedback.

CHAIRMAN APOSTOLAKIS: The microphone.

DR. LOIS: I'm sorry. Early feedback on what we're doing.

For the purpose of brief overview of the human reliability program, I created a picture here that represents the human reliability program as part of the probabilistic risk assessment. I guess very frequently people forget that HRA is part of PRA.

CHAIRMAN APOSTOLAKIS: Excuse me. Can you move to that chair? Because you're blocking the view. Thank you.

MR. YEROKUN: Okay.

DR. LOIS: So when we do a PRA, we start out with identifying plant challenges, initiating events and identify how the plant will respond to those challenges. And as part of that, the system performance and operator actions. And in the PRA actually we describe the possible planned responses and the consequences.

So human reliability is the portion that deals with operator performance of the PRA.

And to perform human reliability we have established a process which starts again with identifying the human actions that are needed as part of the planned response. Decide what is the scope of the analysis, where we should put the actions in our model the event tree, etcetera, how we would deal with dependencies and then to quantify.

And quantification, in order to quantify human actions, we have developed what we call knowledge-base. We have to understand the plan preparedness, plan programs, training decision, etcetera and how those are implemented by the plan as well as we have to understand how people would react under accident conditions or not normal conditions. All that develops what we call knowledge-base and feeds into the various techniques that we're using to quantify.

And if we were dealing with a physical phenomena, ideally we would pick the knowledge-base and use some clear mathematical constructs to describe the phenomena. That's not the case yet in human reliability. And as you can see here, we have several methods that try to depict human performance during accident conditions.

And underneath that I'm going to discuss what are the issues that pertain to each one of these steps. And with respect to the HRA process, we have issues that were talked before the presentation, how well the various steps are performed, when we perform an HRA, consistency among analysts for performing HRA using the same or different methods. And the other constraint we have is that current methods primarily address full power reactor mode and while low power shutdown and external events are also important from a human reliability perspective.

And what do we do about it? We mentioned that EPRI long time ago has developed SHARP 1 establishing the steps for performing human reliability. The ASME developed standards and I guess ANS developing standards for low power shutdown.

The ASME went a level below that and developed the Good Practices to support the standards in limitation for human reliability. But we have to expand those, the guide and development, to new reactors as we develop HRA methods, low pressure down, external events, etcetera.

With respect to the knowledge-base -- I'm sorry, this is kind of --

CHAIRMAN APOSTOLAKIS: It's fine.

DR. LOIS: Taken from one and I guess PC to another changed the fonts, etcetera.

The big issue is understanding human performance under accident conditions. And within that, what are the important performance shaping factors and how the performance shaping factors interact, and what are the dependencies. And again, we have a better knowledge-base developed -- I'm sorry, full power and reactor generation.

We believe that EPRI expanded the knowledge-base, brought in the issue of the errors -- dealt with the errors of commission, identified the importance of contextual aspects on human performance during accident conditions. But we continue to improve. We're collecting data. We have a database where Halden is helping us in developing on performing simulator experiments. And we're starting new work, as Jimy suggested, for new reactors. And hopefully we'll get to low pressure down and external events.

With respect to the techniques, the issues are that none of them appears to have encompassed all of the phenomena that have taken place regarding human performance under accident conditions. There is the issue of consistency of method application and still disagreement among methods and what method is better, what are the important PSFs and how they interact.

In terms of resolution, we did the

evaluation of methods with respect to Good Practices. That is we perceived that this is the first step towards accomplishing a better understanding and agreement among methods.

Currently we focused on domestic methods. In the future we're going to look at the nondomestic methods. We're developing Bayesian tools that would assist configuration. And we plan to use the Halden facilities to test and benchmark the methods eventually.

MEMBER BONACA: Under resolution that we have ATHEANA, where did you have SPAR-H?

DR. LOIS: SPAR-H we'll come to discuss. SPAR-H, we believe that because it is built a lot on ATHEANA, used a lot of the concepts, it has its own entity though. It --

MEMBER BONACA: But it has those performance factors as considerations of that. Now clearly reading the material it's communicated that ATHEANA is a superior method. But it will be interesting to understand how superior. ATHEANA is like a nuclear weapons; it's hidden and is never used. So we are left with big questions about that.

DR. LOIS: And these are the issues of interest that we are going to discuss today and will address.

MEMBER BONACA: Okay.

DR. LOIS: Okay.

MEMBER BONACA: Okay.

DR. LOIS: So with that overview --

CHAIRMAN APOSTOLAKIS: I have a question, Erasmia?

DR. LOIS: Yes.

CHAIRMAN APOSTOLAKIS: Has anybody from NRR ever said in reviewing a licensee application I cannot make a decision here because the human reliability analysis is not good enough or I don't have enough information? Have they ever said that?

DR. LOIS: Yes, they have. We have --

CHAIRMAN APOSTOLAKIS: Because my impression is that they always make a decision.

DR. LOIS: We have a lot of interaction. As a matter of fact, the Good Practices and the evaluation of HRA methods came as a recommendation from NRR. When we did the evaluation of the various PRAs for the purposes of the Reg. Guide 1.200, which is the PRA quality, we were part of the team and evaluated the licensee's HRAs.

So do we have everyday question on HRA? Probably not. But NRR has its own experts, HRA experts if you wish. But we are having a lot of interactions with NRR.

Also it has to be recognized that it is the user if you're in regulatory space is evolutionary in the sense that in the past we were using PRA just for specific purposes and now we're using it in licensing space, etcetera, etcetera. Therefore, the technology, the PRA, the issue of quality of PRA, HRA and how well the various methods are suited for various applications, it becomes more and more apparent and it's needed to be addressed.

MR. YEROKUN: If I may just add, it's also not so much an issue of somebody in NRR coming up with I can't make a decision unless I have HRA input, but it's more I need more input from HRA to make a better decision.

For example, the rulemaking activities. You're familiar with the rulemaking, proposal making for manual action would be heavy HRA involvement in trying to develop support for that. It could be going in a different direction, but there is still the HRA involvement in providing support for whichever way that goes. So it's more we need HRA to make a better, more risk-informed decision as opposed to not being able to make it.

DR. LOIS: I think we'll ask John and Alan to come.

So the first topic to discuss is the evaluation of HRA methods against the Good Practices. Dr. Forester and Dr. Kolaczkowski, both help here. But I'll say we'll explain later. Actually, we have taken the input of the general HRA community.

In terms of background or in terms of outline, I'll discuss the background, why we do this work, what we have done. I'm going to just remind what are the Good Practices or the HRA approaches.

I'll summarize the results and then we'll discuss the individual methods. And at the end we'll talk some of what we learned and where we're going to go next.

Why we do this work? I guess, as we said before, to address PRA quality issues for the use of PRA in regulatory space.

We're developing guidance for performing in reviewing HRA in two phrases; the Good Practices was phase 1, the evaluation of methods against Good Practices is phase 2.

The status is that we have created a draft report which we have for internal review. And that includes the ACRS Subcommittee. We're going to go to full ACRS Committee in February as it's planned now. We plan to publish it for public comment in March and then revise for publication in September.

CHAIRMAN APOSTOLAKIS: So if you get any comments from the ACRS in February, you don't plan to incorporate them before public comment?

DR. LOIS: We'll try to address this. We hope that this discussion with the Subcommittee will give the opportunity to ask to get the bulk of the comments. And going to the full Committee we hope we will have addressed the more crucial ones. But a month in between will be, hopefully, enough. But that's a good point. And probably we should -- it just depends on how many comments. We can always change from March to April.

The approach that we took to evaluate the methods is we started out comparing the methods step-by-step with the Good Practices. And, indeed, we gave ATHEANA and SPAR and SLIM/FLM to external review. Jeff Julius reviewed ATHEANA and SPAR-H and SLIM/FLM.

CHAIRMAN APOSTOLAKIS: So let me understand. When you say "review," you mean their comments are what appear here in the document or that was a separate review.

DR. LOIS: No, no, no. Their comments in

this document. But, however --

CHAIRMAN APOSTOLAKIS: Ah, so the review, the comments we see in the document on ATHEANA and SPAR-H come from outside?

DR. LOIS: And as a matter of fact, what we did, if I finish. We had this initial review. And then we had an expert meeting in June where we presented the results of this initial review. And Jeff was there and Wendall was there, and many other experts. The Idaho HRA group --

DR. FORESTER: People from NASA.

DR. LOIS: People from NASA. The Halden people. We had quite extensive HRA expertise. And we presented the results. And as part of that activity, it was recommended that we should look deeper into the underlying technical basis and address the underlying technical basis as well. Because the Good Practices do not go as deep in the quantification aspect of it.

And then also it was recommended to discuss the methods as intended to be used versus has been used, practiced.

And also we had a session on what is needed, what we should do from now on. And that was also part of the meeting.

So we revised the reviews. And this

revision hasn't been seen by the reviewers of ATHEANA, SPAR-H and SLIM. This is the first time that. We have not communicated with these extended reviewers. I think we will through the public review process.

So we revised the --

CHAIRMAN APOSTOLAKIS: On other thing.

DR. LOIS: Yes.

CHAIRMAN APOSTOLAKIS: When you ask people from the outside to review these models, are you compensating them for their time?

DR. LOIS: Yes. So that was NRC's -- it was not a public review process.

CHAIRMAN APOSTOLAKIS: Okay.

DR. LOIS: It was contractual process through the NRC. But it was, again, with respect to Good Practices.

CHAIRMAN APOSTOLAKIS: Yes.

DR. LOIS: So we have expanded their review to address the underlying one.

And here we are for your reviewing feedback.

I don't think I should focus on that. This applies whether the Good Practices -- remind ourselves what we're going to talk next.

These are the methods that were reviewed.

It was domestic methods, those that are used by licensees and NRC.

CHAIRMAN APOSTOLAKIS: So you didn't feel any need to review MERMOS or CREAM?

DR. LOIS: That will be the next step.

CHAIRMAN APOSTOLAKIS: So you will include them later?

DR. LOIS: Yes. Right now the scope of our work was those methods that are primarily used by licensees for applications and also by the NRC for its own evaluations.

What are the results? The summary? Well, actually, it was recognized that most of what we call methods are just quantification tools. Very few methods provide guidance on how to do human reliability and up to the analyst to decide what are the steps and how well would implement the steps. An exception is ATHEANA that it is provide a method on how to do an HRA.

With respect to guidance on how to do a human reliability, again we mentioned here the EPRI activities. That they do very good job having many of the Good Practices. And since this is an early work, the issue of identifying errors of commission and contextual aspects were not covered.

The HRA methods that are used by EPRI typically are referencing SHARP and SHARP1. But on the basis of NRC's reviews, earlier studies, at least this point here does not have experience on the kinds of applications that EPRI covered this morning. We haven't seen this in production, any of this. But IPs, etcetera, really show question mark whether or not the SHARP and SHARP1 guidance was used as part of the analysis.

With respect to the quantification tools, actually what we see here is the quantification tools are THERP, ASEP, ASME, etcetera. It reflects an evolution of the thinking or an evolution of people's understanding of what are the important inferences on human performance when they respond an initiating event or an accident condition.

Also, early methods are a little bit more simplistic. They address human behavior in a more simplistic manner.

And as methods progress, they become more complicated but also bringing a better understanding of human performance. And also the advances of the social and behavioral sciences that they did through reviewing events and also performing research examining those issues.

And different approaches have different capabilities into capability to translate this qualitative information, the underlying knowledge-base into a number.

Also a note here is that different methods are development and have developed for different purposes.

Some of the strengths. Some methods provide very good and clear technical basis of the underlying method. A good step-by-step guidance on how to use the tool. And also traceable analysis. And it doesn't mean that the same method in those strings are related to different methods.

Weaknesses, weaknesses with respect to the technical basis that some methods are using. And here I make a point that these evaluation appears to lead to indicate that some methods have questionable basis to the point that its use may not be desirable.

CHAIRMAN APOSTOLAKIS: So that was one of the things that I noticed as I was reading the report and we'll come to individual methods later, but let's make a general comment here. The general tone is, you know, you don't go beyond saying questionable, or you might say the validity should be justified. Is that indirect way of saying to people don't use it? And if it is, why don't you just say it or is it too soon to say that? Because you're putting a tremendous burden on the reviewer who presumably will use this document. The poor guy, you know, doesn't know what you know. And he sees here words like -- I'll tell you in a second. "The validity of such generalizations is questionable. There will be a great deal of uncertainty in the results obtained using these method." And then there's a whole list of weaknesses and at the end there are five lines that say, on the other hand there are some strengths.

You are indirectly telling the world it's better not to use this method. I'm wondering why don't you come out and say that?

DR. LOIS: In the meeting we have the expert meeting that we had in June discussing all of this, we were debating whether or not we should say this method is very weak and therefore not applicable or should not be used. On the other hand, people felt that methods may be good enough for some applications and therefore if you do a very high, you know, a conservative analysis or a high level analysis, maybe ASEP may be okay. For a more detailed analysis may not be.

So the concept of the tool bags was kind of more recommended as opposed to totally disregarding methods. However --

CHAIRMAN APOSTOLAKIS: You know, this perpetuating the situation where we have a bunch of models out there.

Anyway, go ahead.

DR. LOIS: However, I think we're kind of willing to identify some of the methods that may be more -- less desirable to be used. And also the next step that we believe that should be taken is do a Reg. Guide or an SOP which characterizes the capabilities of the method for what application. And that clarify further.

CHAIRMAN APOSTOLAKIS: Yes. I mean, I appreciate the difficulty of generalizing and saying, you know, you will recommend yes, no on every method. No, you can't do that because some methods indeed may be useful in some instances. But in a case where the whole thing rests on some very questionable assumptions, it seems to me you should send a clear message that the NRC would not be willing to entertain, you know, applications that involve this method. Because this happens in every field that is new, although I don't know how new this is, but it's new, it doesn't have an established state of knowledge and so on. But there are all sorts of models and methods and people are reluctant to express strong views. But eventually we know that some of these methods will sink.

And this reminds me of the PRA procedures guide of 25 years ago when people were not sure whether Bayesian methods were the right way to go, there were vested interests and so on. So it says here's one way, here's another way. And then what do we see years later? No one's using.

So I think in some cases you have reached the point where you can say -- you know, you don't have to say this is stupid, but you can say it is not advisable to use this method or something to that effect. I think that would be much more useful to the reviewer.

Because remember, the reviewers they have other things. They have to approve a licensee application and so on. They cannot go back and red the whole literature to figure out. And when you tell the reviewer the use of this method is questionable, I don't know what he or she can do with that.

So that's something that I think, you know, is something you want to consider.

DR. LOIS: Absolutely. And I think that's my last bullet we tend to go towards that.

CHAIRMAN APOSTOLAKIS: Okay.

DR. LOIS: But your input is very valuable here.

CHAIRMAN APOSTOLAKIS: So you can put at the end this method is dropped.

DR. LOIS: Recommend not to be used.

CHAIRMAN APOSTOLAKIS: Ah.

DR. FORESTER: Yes. I guess I would comment. In some cases there may be some data out there that is proprietary or something that, say, we can't really make the final decision necessarily. It just appears to be that way.

CHAIRMAN APOSTOLAKIS: If it is proprietary, John, you reject it. If you don't have access to the basis of the method, you say the NRC will not review applications of this.

MR. KOLACZKOWSKI: This is Alan Kolaczkowski.

And my only comment, George, is that now that's an NRC policy decision. As NRC contractors, we can perhaps provide some advice to the NRC, but that's an NRC policy decision.

CHAIRMAN APOSTOLAKIS: It is a policy of the Agency. I mean, we are not approving results of methods we have not reviewed, right? So, you know, why should HRA be any different?

DR. LOIS: And the word "review" here should be qualified because it's more with respect to Good Practices. It's not a review in the --

CHAIRMAN APOSTOLAKIS: It's up to you experts to decide. I mean, I'm not taking any latitude you have.

DR. LOIS: The word review, that is review--

CHAIRMAN APOSTOLAKIS: But I mean the last several years I have seen detailed reviews from the staff on Westinghouse reports, General Electric reports and they're all proprietary but the staff has reviewed them. The staff is comfortable. They have made comments. GE came back and said this is how we respond and so on.

Okay. Findings?

DR. LOIS: With this broad overview, what we're going to discuss here, John and Alan, the scope of the methods, the underlying model data, quantification approach, strengths and weaknesses; that's how the presentation has been structured.

CHAIRMAN APOSTOLAKIS: Good.

DR. LOIS: Who's going first.

DR. FORESTER: Alan's first.

CHAIRMAN APOSTOLAKIS: So you're going to go over all of them?

DR. LOIS: Yes.

MR. KOLACZKOWSKI: Yes, but giving them in total we're going to try to save some time. What we'll all do, as Erasmia just point it out, each review method has a scope slide and then an underlying basis slide, quantification slide and then strengths and weaknesses. I don't think we need to tell the ACRS Subcommittee, remind them what THERP is and what ASEP is, etcetera. So I'll try to go through in each case the scope, underlying basis, etcetera very quickly because I think what's probably more of interest in this presentation is our view of the strengths and limitations.

CHAIRMAN APOSTOLAKIS: But let me ask you this, Alan. Yes, I agree with you.

Look at this bullet that says "Diagnosis contribution to error is handed with time reliability curves?

MR. KOLACZKOWSKI: Yes.

CHAIRMAN APOSTOLAKIS: This is a statement of fact.

MR. KOLACZKOWSKI: Yes.

CHAIRMAN APOSTOLAKIS: Are you giving now any advice to the user what that means? Is that good for some screening purposes or some quick analysis but not so good if you -- I mean, if there is a human action somewhere that is really critical, are you saying you shouldn't do this, you should go to something more detailed?

MR. KOLACZKOWSKI: In the draft report that we have I think we've gone, perhaps part way at addressing your issue. Perhaps we haven't gone far enough.

You'll recall at the end of each review there's a sort of a list of questions that says if you as a reviewer have a submittal and they've done it using THERP, here's some things to think about. And to pick on that one in particular, I believe under some of these methods we've indicated clearly if there's reason to believe that the operator action is dependent not so much on time, it's more dependent on other PSFs, if you will, well then you have to at least question whether just use in a time reliability curve is even the right method to use. Because if you believe it's not driven by time, it's driven by something else, some procedural deficiency perhaps or some environment; he's got to go out in the snow and go turn this valve. Maybe the ergonomics is much more important factor and yet you're pretending to believe that the diagnoses is driven by time and you're using a time reliability curve. You certainly have to question whether that's even the appropriate method to use.

CHAIRMAN APOSTOLAKIS: But I haven't seen such crisp statements in the report, is what I'm saying. And also you seem to seem to rely on the verb "question" a lot which, you know, the reviewer may not find very useful.

MR. KOLACZKOWSKI: Understand.

CHAIRMAN APOSTOLAKIS: But if you tell him, you know, because of all these reasons in this case don't do this, then I think people understand that. That's all I'm saying. I mean, your recommendations would benefit from a little stronger statements.

MR. KOLACZKOWSKI: Understood. Understood.

CHAIRMAN APOSTOLAKIS: Yes.

MR. KOLACZKOWSKI: Understood.

CHAIRMAN APOSTOLAKIS: Okay. Good. Let's move on.

MR. KOLACZKOWSKI: Okay.

DR. LOIS: Oh, I'm sorry.

CHAIRMAN APOSTOLAKIS: No, that's good. That's good. Never be sorry.

MR. KOLACZKOWSKI: THERP, you know, primarily addresses pre and post-initiates. It's been around for a long time, etcetera. Primarily it breaks human error down into a diagnostic phase and then an implementation phase.

Wait, I want to get caught up here where I am here. Just bear with me. Okay.

CHAIRMAN APOSTOLAKIS: Slide 14.

MR. KOLACZKOWSKI: Okay. And primarily you come up with a diagnosis probability, you come up with an implementation failure probability and then you sum them up to get the total. And it does provide some guidance on assigning uncertainty, the distribution about the number that you get. But that uncertainty distribution, as has already been commented during our earlier presentations, is primarily based on what value you get out of this process.

If you have a .1 failure probability, then it's going to tell you to assign a -- excuse me. An uncertainty bound of more bigger than maybe a factor of five because you don't want the maximum to go greater than one. And on the other hand, if the failure probability is small, the believe is that the uncertainty is larger and it will tell you to assign a larger uncertainty.

The uncertainty doesn't really come from the analysis and the context, etcetera. It's just an assigned value based on whatever point estimate you come up with.

Okay. Next slide.

I've already indicated it primary uses a time reliability curve.

No, let's go on to the next one. I've already covered this.

So what are some of the strengths and weaknesses of the THERP analysis? Clearly, one of the strengths in THERP is especially we're dealing with the implementation phase of the error. It prescribes a rather detailed task analysis so that you really understand what the operator has to do to implement this action, whether it's calibrating a device or whether it's a post-initiator action. And that's very valuable, provides very valuable qualitative insights.

It's been applied widely across many industries. There's a large pool of experienced analysts. A lot of people, for the most part, understand THERP and generally how to use it. It's been around a long time. There's a lot of experience had there, which in a way gives it a strength.

There's a good qualitative discussion of a broad range of potentially relevant PSFs. On the other hand if you look over on the weakness side and particularly the last bullet, unfortunately only a small subset of those are actually they tell you how to treat them quantitatively in the analysis.

So if an analyst wants to treat some of the other PSFs, there's no direct way to do it in the guidance that's provided in 1278, so hence the analyst has to decide how to factor these other PSFs. Like, well maybe I should increase stress by something higher because of some other PSFs I'm looking at. And that's when you start getting analyst-to-analyst variability.

CHAIRMAN APOSTOLAKIS: I'm intrigued by your second bullet under weaknesses. Not implemented as intended.

MR. KOLACZKOWSKI: Well --

CHAIRMAN APOSTOLAKIS: What do you mean?

MR. KOLACZKOWSKI: And again, I think we just wanted to highlight. That again because this has been around a long time and we do have an experience base growing on how people use THERP, unfortunately a lot of people just go into the tables and use the numbers without having read the first ten chapters of THERP so that they really understand how to use those tables and when to pick the right value out of this table or this table or this table. They think they can just go into the table, see the heading, and say this is for pre-initiator umpty-umph and my stress is high, so the number must be .03 and you go use it.

CHAIRMAN APOSTOLAKIS: Now we were told earlier by EPRI that there is a lot of leg work that you have to do before you use. Is there anything there that says go read the first ten chapters?

DR. ELAWAR: Those ten chapters --

CHAIRMAN APOSTOLAKIS: I think there are 20.

MR. KOLACZKOWSKI: Yes, whatever. Seventeen or whatever.

DR. ELAWAR: Are usually read in order to make a decision as to where would I go and which table I would use in the THERP.

CHAIRMAN APOSTOLAKIS: I'm not talking about you personally.

DR. ELAWAR: Well, as far as I know most HRA models do have a lot of leg work in determining where should I go, what should I use. And they're not repeated each time, too.

CHAIRMAN APOSTOLAKIS: Now I wonder whether the word "weakness" is the appropriate word here. I mean, is it really the fault of Swain and Gutman that people don't use it as intended?

MR. KOLACZKOWSKI: No, that's a valid point. Some of the things that are listed in the weakness column are not always a weakness of the method, per se.

CHAIRMAN APOSTOLAKIS: It's a practice.

MR. KOLACZKOWSKI: But it's also a weakness of a common practice that we tend to see out there.

CHAIRMAN APOSTOLAKIS: I wonder whether there's another word that's more appropriate.

MR. KOLACZKOWSKI: Perhaps there is. We could think of something. Negatives and positives about the use of the method are something. Okay.

So that's sort of the story on THERP.

Moving to ASEP. Again, I think most people here are probably pretty familiar with ASEP, so we won't go over the scope and an underlying basis in too much detail.

It's basically a simplified THERP. It was put together so that systems engineers or PRA analysts with perhaps not a lot of HRA background could at least have a method that they could use where they didn't have to read the first 19 chapters of THERP and could still get out what was believed to be a reasonably yet probably conservative number based on a few things to be considered to come up with this HEP.

Its basic approach is to take the pre-initiators, assign a generic error rate and then based on how many checking type recoveries you have, you assign some additional probabilities which tend to lower the basic error rate.

Post-initiators, again just like THERP uses a diagnostic implementation model approach. However, it's a simplified version of both of those models that are used in THERP, but it essentially follows the same process.

Next slide.

I've already mentioned pre and post-initiators are quantified based on an adjustment of essentially a generic or, if you will, in the case of the post-initiators an initial error that you assign. And then you adjust those based on a few PSFs.

I've already mentioned the use of the diagnosis is the same, more or less, as THERP.

Again, there's a fixed set of PSFs. There's limited guidance for how to apply them. You basically go to a series of look up tables and curves and you pick out a number. If in your judgment the stress is high, and it says if you think the stress is high, take the basic HEP and multiple it by five or whatever.

Again, the uncertainty bounds are assigned in ASEP, much the same way as THERP. It's really more dependent on what the value is, not so much what the context is.

Strengths and weaknesses. Easy to use, simplified technique.

Tends to lead to a thorough analysis pre-initiators. A lot of effort went into how to analyze pre-initiators in ASEP. We didn't have that before. And actually does, I think in my people's judgment, a pretty good job of coming up with pre-initiator HEPs.

It does explicitly handle, again, diagnoses and implementation. That's a strength.

And I think, and again this is more of a judgment thing, but I believe the results are commonly accepted as reasonable for what we call not far from average context.

And another positive is that the screening approach does require some analysis. You do have to do some amount of leg work, thinking, etcetera to even come up with the screening values. And that's probably a good thing. At least it forces the analyst to do some thinking, even in assigning screening value.

CHAIRMAN APOSTOLAKIS: What does "average context" mean? Does it mean what most people would anticipate or --

MR. KOLACZKOWSKI: I'm going to put my ATHEANA hat on here now for a moment.

Basically that the scenario is one that operators are used to seeing in a simulator, etcetera, and things aren't so -- like the plant isn't getting into a physical regime that's really almost unexpected, not well understood, etcetera. Now you're starting to get into error forcing context, and that's a whole other issue.

CHAIRMAN APOSTOLAKIS: Okay.

MR. KOLACZKOWSKI: On the weakness side. And, again, this is probably not -- the first one is not so much fault of ASEP, it's just because it is so easy to use, analysts may use the technique without really having the HRA background to use it. It's so easy, it's easy for an engineer with very little HRA background to go in and start picking numbers out of tables and perhaps misapplying it.

Judgments about the PSFs and the context are made by the analysts, again with little guidance. That's why we would almost argue you should have somebody with some HRA background even using ASEP.

It cannot directly handle more extreme or unique PSFs, as I pointed out. It's really good for average context, if you will.

Same data limitations as THERP. All this data is coming primarily from judgment, etcetera.

Next slide.

I'll hand off to John, he's going to cover a few others. And then I'll come back to a few others.

DR. FORESTER: Okay. I'm going to discuss now the HCR/ORE method that was published in EPRI TR-100259 which was mentioned this morning. This is one of the methods that is included in the HRA Calculator.

The method focuses on really on estimating nonresponse probability of post-initiator human actions only.

CHAIRMAN APOSTOLAKIS: Excuse me. Is this the first time that you gentlemen see this, this evaluation? You have not seen it?

MR. JULIUS: I participated

in a meeting, so I saw this.

DR. FORESTER: Yes, Jeff has seen some of this.

So in general, the approach doesn't really address errors, per se. They're just looking for the likelihood of nonresponse. Essentially the assumption is that over time they'll figure things out and they will make a response. So there's not really a focus on errors or they sort of assume the correct diagnoses.

CHAIRMAN APOSTOLAKIS: But speaking of that, I remember reading a paper on the cognitive psychology literature many, many years ago that said that they have done some experiments and their conclusion was that if the subjects had not figured out what was going on within 80 minutes, then they would never figure it out.

DR. FORESTER: Without 80 minutes?

CHAIRMAN APOSTOLAKIS: Eighty, eight-zero. Now, it could have been 60, but I think it was 80. But it's interesting because it gives a different spin to this that, you know, there is a certain amount of time within which people can figure out what's going on. But given a very long time, it's not clear. Well, I think if you give them five years, they'll probably figure it out.

DR. FORESTER: Yes.

CHAIRMAN APOSTOLAKIS: But we're talking about, you know --

DR. FORESTER: Sure, I understand.

CHAIRMAN APOSTOLAKIS: -- giving them three hour versus an hour and a half. That they found that it was irrelevant. I mean, if they couldn't figure it out, they just couldn't. And I'm wondering how relevant that this or whether such a conclusion is supported by other people's experiments. Because that was a single paper.

DR. FORESTER: Right.

CHAIRMAN APOSTOLAKIS: Are you familiar? I mean, you're a psychologist?

DR. FORESTER: Yes, I am. I'm not familiar with that paper, per se. But, you know, generally the kind of time frames we're looking at in accident scenarios move a little faster than that. And they will be forced to do something eventually, fairly quickly generally.

CHAIRMAN APOSTOLAKIS: But will they figure out what's going on; that's the question.

MR. KOLACZKOWSKI: George, this is Alan Kolaczkowski.

Again, it's an event a long time ago, but TMI, I mean look, it went for quite a few hours. And they didn't really understand what was going on before that operator came in on a shift change and said, you know, I think we may have the PORV stuck open. And that was many hours later. And then they finally closed and get an injection going, etcetera.

So sometimes new cues, new person, whatever all of a sudden it's a whole new ballgame.

CHAIRMAN APOSTOLAKIS: Yes. But if a model, though, puts a distribution there that has a pay --

MR. KOLACZKOWSKI: You could still maybe do it.

CHAIRMAN APOSTOLAKIS: You know, maybe within some reasonable time you figure it out, then you have to question that, right?

MR. KOLACZKOWSKI: Yes.

CHAIRMAN APOSTOLAKIS: But I'm not sure that any of the models consider saying anything. Maybe when we talk about Halden, maybe they can figure out an experiment to see whether that is a valid thing?

DR. LOIS: They're doing some experiments.

CHAIRMAN APOSTOLAKIS: On that subject? I know they're doing experiments.

DR. LOIS: No. But we can --

MR. KOLACZKOWSKI: In ATHEANA in the recovery step one of the things you consider is what are new cues available, is there new staff available, etcetera, or could you be in a mindset that therefore you're just never going to figure it out.

CHAIRMAN APOSTOLAKIS: Yes.

MR. KOLACZKOWSKI: So I think it's somewhat addressed in there.

CHAIRMAN APOSTOLAKIS: Well, let's see if they can figure out an experiment.

DR. FORESTER: Okay. Yes, there's cognitive aspects like tunnel vision where people get focused in on a particular kind of diagnoses and there's anxiety involved and so forth and they will tend to focus. But as Alan pointed out, sometimes other cues will come up later on that may get them -- it's certainly possible.

DR. LOIS: Let me rephrase. In some of the Halden experiments time has been used as a measure of success or completion of the task, etcetera. So we'll have some information later on on that.

DR. FORESTER: I'll just note, too, that the HCR/ORE method as written in that document does include the CBDT method, too, to address the longer time frame events. So the ACOREs, the TRC, which I'll talk about, for long time frame events, the CBDT is recommended.

I do need to give you a little bit to understand the underlying model I think for this method so we can discuss the strength and weaknesses.

As it indicates there, it's a simulator measurement-based TRC. It relies on a couple of parameters, of estimating a couple of parameters. And this can be obtained from crew response data. They look for the meeting response time in a particular accident scenario and the standard deviation, so they look for a measure of variance.

Then the idea is that if you have those parameters, you can estimate the probability of nonresponse within a given time frame using the standardizing normal committed distribution. So the basic idea is if you know what the median response time is, you have an idea about the standard deviation, you can essentially look up the probability in a Z table.

Now, the basic approach is really based on a series of experiments that were conducted by EPRI called the ORE experiments, operator reliability experiments. And the idea was that they would go to several different plants and they'd run different crews through different kind of accident scenarios. And they'd look for how long it took them to respond. So they'd get an estimate of the median response time and therefore, then could derive an idea about the variance and standard deviation. And that this then generic information that was obtained from these experiments looking at the both the crews and both the PWRs and BWRs, that then this generic data could be used by other licensees for their IPEs and so forth. So that was the basic idea, was to get that kind of information to support that process.

CHAIRMAN APOSTOLAKIS: They give you one value for the median and one value for sigma? But they don't give you any uncertainty about this? Is that true?

DR. FORESTER: It's true. Yes. I guess another goal of the method was also ACR was a sort of proceeding methodology and there was some assumptions in ACR that they wanted to test. So that was another reason for doing the ORE experiments.

CHAIRMAN APOSTOLAKIS: Speaking of the equation, by the way, there's a typo on page 57. You have caught it? The equation is not correct?

DR. FORESTER: It's not correct. No, I

haven't caught it then. I'll get it from you.

CHAIRMAN APOSTOLAKIS: Thank you.

DR. FORESTER: Thank you.

CHAIRMAN APOSTOLAKIS: Nobody's asking me what it is. Okay. Let's go.

DR. FORESTER: Oh, I want to know, but we can do it later if you want.

CHAIRMAN APOSTOLAKIS: All right.

DR. FORESTER: Okay. Given that approach, in doing the experiments they sort of realized that there are plant-specific differences. So ideally, it's probably not a good idea to use the generic data to take the data from their experiments and use those for another entirely different plant.

CHAIRMAN APOSTOLAKIS: If they give you on several things, they might say that my plant is here or there. But if it's a single point value, that makes it even more difficult.

So they tell you to go to expert judgment?

DR. FORESTER: Essentially, yes. Well what they ideally they want you to do if you want to use the approach for your plant, you would identify the human events you want to quantify and the relevant accident scenarios and you would run your own crews through those scenarios and get your own estimates and median response time. Then you could calculate the standard deviation. That would be the ideal approach. Of course, that's going to require running a lot of crews through a lot of simulations, which we'll get back to later.

If that's not available, another recommended approach for obtaining the parameters is to just use expert judgment from operators. So basically they would ask the operators how long they think it would take them to respond in this particular kind of a scenario.

They do have some ideas about you might use the calculations to let them know when certain parameters would be available and so forth. And then from that, they would be able to try and make those judgments.

CHAIRMAN APOSTOLAKIS: But I don't remember in the document that you are actually commenting on this, that operators may be optimistic. Are you saying anything about it?

DR. FORESTER: What we're focusing on it is that it's questionable because --

MR. KOLACZKOWSKI: It's questionable, yes.

DR. FORESTER: -- there's no guidance given for how to do that.

CHAIRMAN APOSTOLAKIS: You certainly comment about that, yes. You would like to see guidance. Ah, okay. You do have a sentence, aside from the concerns about operators being able to make estimates of when they would be likely to do something, the method provides very little guidance. Yes. But this is an important issue, though. And I think I read another paper a long time ago that stated really the obvious, but they had evidence, that the operators tend to under estimate the time it will take them to do something.

DR. FORESTER: That's true. And that's one of Swain's -- actually, that was mentioned in Swain, too. Anytime you use an estimate from an operator, his recommendation is double it.

CHAIRMAN APOSTOLAKIS: I mean, there is no implication here that there is malicious attempt on their part to achieve.

DR. FORESTER: No, no, no, no.

CHAIRMAN APOSTOLAKIS: They truly believe this.

DR. FORESTER: That's true.

CHAIRMAN APOSTOLAKIS: Which is a standard example of over confidence, I think. People are more confident than they should be.

DR. FORESTER: Yes. In my experience I haven't really seen any cases described where expert judgment was used, but there may be some out there.

CHAIRMAN APOSTOLAKIS: There is another interesting statement you have here. The potential for an actual diagnosis error and the resulting effects of an incorrect response are not explicitly addressed in the HCR/ORE method. What was that mean? I mean, they will tell you they calculate the probability of nonresponse --

DR. FORESTER: Right.

CHAIRMAN APOSTOLAKIS: Are you saying what if they take the wrong response, what happens, is that what you mean by this?

DR. FORESTER: Well, that's one thing. What happens is if they fail to make a diagnosis. Basically, this method by just looking at nonresponse probability, they're sort of assuming that diagnosis will occur and will be correct. But there is a possibility that errors will be made in the diagnosis and that an inappropriate action could be taken.

CHAIRMAN APOSTOLAKIS: Okay.

DR. FORESTER: And that really isn't addressed.

CHAIRMAN APOSTOLAKIS: It's not addressed.

How about if they tell you that it's not the business of HCR to do that? It's the business of the PRA analyst who develops the event tree so that you will have a different branch that says, you know, wrong diagnosis and you do something else? You know, it depends on what the method is intended to do. I don't think they're going to tell you that, but they could.

DR. FORESTER: They could.

CHAIRMAN APOSTOLAKIS: In fact, now they might.

DR. FORESTER: They might.

MR. KOLACZKOWSKI: And I think that's what we're trying to indicate here. And I know you want us to make stronger statements in the report. But if a submittal comes into the NRC and they've done, in this case let's just say HCR/ORE and no other method or something, you have to recognize it doesn't treat diagnostic failure probabilities. And so if the reviewer believes that this situation is so complex that maybe the operator wouldn't even recognize what is the right action to take, well then you got to recognize that the method doesn't treat this. So hopefully the submittal has already treated the diagnostic part of the concern, if there is one, with one of the other methods and now the combined answer is really the total answer.

So we're trying to indicate to the reviewers what are the weaknesses, perhaps there's a better word. What is in the scope that you need to recognize that this treats this but doesn't treat this. This does this very well, this does this not very well so when a submittal comes the reviewer understands what the scope limitations are, what the weaknesses are even in the stuff that it does treat, etcetera. And then look at that submittal with those eyeglasses on.

CHAIRMAN APOSTOLAKIS: Speaking of that, I just remembered. I thought one of the good steps forward in the development of human reliability analysis was, I think they called it confusion matrix about 20 years. Where it was a matrix with initiating events.

MR. KOLACZKOWSKI: Yes.

CHAIRMAN APOSTOLAKIS: And the idea was to show that the symptoms of this event might lead the operators to think that something else has happened. And in a lot of the cases, in fact they concluded that even if the operators misdiagnosed, they would take actions that would be beneficial anyway. I didn't see anything on the confusion matrix anywhere. Is anybody using it? I thought it was a pretty good thing, or HRA comes after that?

DR. FORESTER: No. I think there may be some people using it. There's a couple of papers in the late '80s, I think, where they --

CHAIRMAN APOSTOLAKIS: No. But the methods that are being reviewed here --

DR. FORESTER: Well, we haven't reviewed that as a method. I mean, that's almost a tool that you'd use with any even method, possibly. It might be a tool that ATHEANA might use. It might be a tool that other methods would use.

CHAIRMAN APOSTOLAKIS: But shouldn't that be part of the discussion that the issue of confusion and misdiagnosis is not as bad as we originally thought and here is some evidence that, you know, that people have thought about it. It was really a very good paper that was published. I don't remember who wrote it.

DR. FORESTER: Who was it?

DR. COOPER: It was Gordon who wrote it?

CHAIRMAN APOSTOLAKIS: Gordon?

DR. COOPER: Yes.

CHAIRMAN APOSTOLAKIS: But a lot of people really felt relieved because before that there was a diagnosis, on my God, we're in trouble. And then the guy comes in and shows you that it's not a big deal. It's really not a big deal.

I think the method should put that somewhere there. And I don't know whether your report should do that, but I thought maybe you should say something about it, I don't know.

It's not a method, you're right. It's not a method. It's just a step in developing naturally the event tree; that's really what is it.

DR. FORESTER: Your point was that even in a lot of cases in power plants, for example, even though they may diagnosis it --

CHAIRMAN APOSTOLAKIS: That's right.

DR. FORESTER: -- the responses may still work out.

CHAIRMAN APOSTOLAKIS: The response still works out, which is really a very comforting thing to know.

DR. FORESTER: Yes, that's true.

MR. KOLACZKOWSKI: Hard to do bad thing.

DR. FORESTER: Okay. I guess one final thing I want to point out here is that by doing this kind of -- just looking at performance in simulators, really there is no attempt to identify PSFs or factors that might create problems or plant conditions that might create problems. It really is a more simple approach than that.

CHAIRMAN APOSTOLAKIS: Are you done with this?

DR. FORESTER: Unless you want to talk about strength and weaknesses.

CHAIRMAN APOSTOLAKIS: In the report there are a couple of things I want to mention.

DR. FORESTER: Okay.

CHAIRMAN APOSTOLAKIS: On page 64 there's while this conclusion may very well be the case, the data on which it is based is proprietary and not available. Now that's three red flags for me. It's not available to you. If it's not, I would say don't use it.

DR. FORESTER: Well, I will say --

CHAIRMAN APOSTOLAKIS: I wouldn't hesitate.

DR. FORESTER: I will say that I have asked EPRI for other kinds of information, and they've been very helpful with that. Yes, that's right, because that's the real detailed data from the ORE experiments, and we do not have that.

CHAIRMAN APOSTOLAKIS: But shouldn't you have it? I mean, if the whole method is based on those, you should have access to them and treat them, you know, with appropriate care.

DR. FORESTER: Well, what that is is the basis for using the underlying distribution.

CHAIRMAN APOSTOLAKIS: That's a big deal here, isn't it?

DR. FORESTER: But even beyond that --

CHAIRMAN APOSTOLAKIS: And then General Physics Corporations also did experiments and you say why the validity of this data is unknown. I mean, how can you use works like that in a regulatory space? You can't. It can't be unknown to you.

And then another comment. There is a paragraph here that makes absolutely no sense to me, but maybe it does and you guys can go and correct the presentations. Page 64, the last full paragraph. It talks about two screening approaches that are suggested in TR-100259. I have no idea what you're saying here.

DR. FORESTER: Page 64?

CHAIRMAN APOSTOLAKIS: Yes. Does the method allow for the use of screening conservative values particularly during initial evaluations of

HEPs? And you say yes. And what follows the yes is incomprehensible. You don't have to explain it now.

DR. FORESTER: Okay. But make it comprehensible?

CHAIRMAN APOSTOLAKIS: Yes.

MR. KOLACZKOWSKI: So noted.

DR. FORESTER: So noted. Okay.

CHAIRMAN APOSTOLAKIS: The rest of it, by the way, reads very well. I mean, I think it's a very impressive document. This is very good.

MEMBER KRESS: What's the error on page 57?

DR. LOIS: That was page 64, nothing else, right?

CHAIRMAN APOSTOLAKIS: The brackets after the F. And if you're familiar with Word, by the way, the brackets can be bigger than they were.

DR. FORESTER: That was the problem. I don't need it after what?

CHAIRMAN APOSTOLAKIS: After F. F brackets dot.

MEMBER KRESS: Dr. Apostolakis, I am very impressed. You read this in detail, didn't you?

CHAIRMAN APOSTOLAKIS: Yes. Because I knew you would be here. I knew you would be here and

I had to.

So I want to make sure that everybody understand that I really like this report.

DR. FORESTER: Great.

CHAIRMAN APOSTOLAKIS: These are comments to improve it.

DR. FORESTER: We hope the general public will feel the same way.

CHAIRMAN APOSTOLAKIS: General public?

DR. FORESTER: The licensees, EPRI, etcetera, etcetera.

CHAIRMAN APOSTOLAKIS: Ah, you guys know you were the general public?

Okay. Yes, I already said. I mean, when you tell the guys the reviewers given the potential impact of the variation and the sequences, the validity of such generalization is questionable, there will be a great deal of uncertainty in the results and so on, you're essentially telling them, you know, this is not very good but you don't come out and say it. And at the very end, you felt that you were too critical. So you say there are some strengths to this method.

DR. FORESTER: Well, you know there is strengths --

CHAIRMAN APOSTOLAKIS: This is the weak side of you.

DR. FORESTER: Okay.

CHAIRMAN APOSTOLAKIS: I mean, after a long list of like two pages of bad things. You say, you know, it may be all right.

DR. FORESTER: Well, we do really like to see lots of simulator exercises. To the extent that they're willing --

CHAIRMAN APOSTOLAKIS: Right.

DR. FORESTER: -- to do a whole lot of that kind of work, that's good information.

CHAIRMAN APOSTOLAKIS: Yes. No, I agree. And I also would like to see them. Don't say they are unknown. You know, if you'd see them, we'd all be happy.

Okay. Are you done with this method?

DR. FORESTER: Yes.

DR. FORESTER: The next is the CBDT, which is also part of TR-100259. Again, it was develop to deal with the longer time frame scenarios where time may not be an issue to avoid optimism.

And this, as I said, it was developed in that context but I think over the years CBDT has become to use more stand alone type of method. And I think even within the HRA Calculator it's indicated as being used. It's a default method rather than --

CHAIRMAN APOSTOLAKIS: Let me ask you something else that has bothered me for years.

DR. FORESTER: Okay.

CHAIRMAN APOSTOLAKIS: In your review of these methods have the developers of any of these methods said anywhere and we are using the results of this other guy and we're building with it, or is everyone starting from scratch?

DR. FORESTER: At that period of time there's a lot of starting from scratch, except that most of these methods do rely on the data that was contained within THERP to adapt that data to do the quantification within the newer method. But in terms of how they go about it, it's usually very different.

CHAIRMAN APOSTOLAKIS: Because every time I see a report or a paper from this community it appears that they're working in a vacuum.

DR. FORESTER: Well --

MEMBER KRESS: And in reality, they're not. But perhaps it's not enough of an official recognition or whatever. I mean, to the extent a method is treating human error as a diagnostic and an implementation phase, I mean you can trace that back to THERP.

CHAIRMAN APOSTOLAKIS: Sure.

MEMBER KRESS: And even prior to that time. Do they actually acknowledge that officially in their report? Many times we don't. I don't know why, but we don't.

CHAIRMAN APOSTOLAKIS: Even they're 35 to 55 references at the end, it's not clear how they really do it. Maybe, you know, it's time to start doing that --

MEMBER KRESS: You can tell there's has been an evolutionary process.

CHAIRMAN APOSTOLAKIS: Right. That leads to another question that I had about the document itself.

DR. FORESTER: Yes.

CHAIRMAN APOSTOLAKIS: There is a review as you are presenting here of the various methods and models which is, for example, let's say. Wouldn't it be nice to say somewhere if it's appropriate that a particular method is more general and it includes all the useful things that two other methods have? In other words, have some sort of maybe hierarchy and say, you know, if you go with ATHEANA for example, then all the stuff is included in the context and this and that; it's more general of other methods. If you go with this CBDT, it includes the good things of HCR, but some of the bad things, perhaps, it includes other things.

I don't know whether that's feasible, especially with the time pressure you have on you now. But I think a user would probably find that useful too to say well gee, okay, they're telling me that this is questionable but then it has some good things. But if I go to this other method, then I'm covered.

I don't know. Is that feasible?

DR. LOIS: However, it may be feasible, but we do view the methods maybe more applicable, various methods more applicable for different applications.

So, for example, ASEP was created because of the extensiveness of THERP and the time needed, etcetera.

CHAIRMAN APOSTOLAKIS: Right.

DR. LOIS: So if you do the current here, ASEP is second to THERP. But --

CHAIRMAN APOSTOLAKIS: And you can say that. You can say that THERP is more detailed, but ASEP has certain -- well, the problem -- I mean, the problem, it's not a problem. But ASEP and THERP you probably don't have that issue there because they were developed by the same guys, right? It was Swain essentially behind those methods. So I'm sure in ASEP he says, you know, I'm using a lot from THERP.

DR. FORESTER: Yes.

CHAIRMAN APOSTOLAKIS: But when you have a separate group developing a method, then you know that they're relying on somebody else but they don't say -- and their method, perhaps, is broader, than it would be helpful to -- if there are such insights. If there aren't, you don't do it. I mean, it's not that you have to try to desperately to do it.

DR. FORESTER: I understand. It's worth thinking about, though. To structure something like that, sure.

CHAIRMAN APOSTOLAKIS: Yes. So, gee, you're so slow, John.

DR. FORESTER: I know. That was a hard one, though.

I mean, we don't have to spend anymore time than you guys want to on this.

The CBDT, again, it's a little bit unique in the sense for that time it did begin to focus on causes of human errors.

CHAIRMAN APOSTOLAKIS: So let me ask this then along the same lines as the previous comment. Has EPRI tried to remedy some of the weaknesses of HCR in the CBDT method? That would be a useful insight. Your question did a lot to the HCR.

DR. FORESTER: Yes.

CHAIRMAN APOSTOLAKIS: So if I go now to this more recent model, are some of these questions removed?

MR. JULIUS: The CBDT model was developed as a follow-on to the HCR looking at the limitation of the HCR/ORE. And that was the reason for developing the cause-based decision tree model.

CHAIRMAN APOSTOLAKIS: Right. But that doesn't tell me whether you have removed some of the questionable part of HCR. Are you saying that it's really HCR but more up to date?

MR. JULIUS: It did remove by breaking out or modeling explicitly some of the casual factors causing you to look at things that were implicitly included in the timing in HCR.

CHAIRMAN APOSTOLAKIS: But the fundamental equation of the log normal is still there?

MR. JULIUS: For HCR.

CHAIRMAN APOSTOLAKIS: You see --

MR. JULIUS: No, no. We go away completely to eliminate that equation and go to decision trees.

DR. FORESTER: This is an entirely different approach.

MEMBER KRESS: An entirely different approach, George.

DR. FORESTER: It could stand alone. It doesn't rely on HCR/ORE at all unless you have a short time frame, then it's unclear exactly how you would deal with it without going to some method. Because the CBDT itself does not address time, shorter time frame than that.

CHAIRMAN APOSTOLAKIS: It says here it serves as a check on cases where the HCR has produced low values.

DR. FORESTER: That was it's intent.

CHAIRMAN APOSTOLAKIS: Does it mean that I do HCR first and if I find low values, we'll do this. Or that was the original motivation for EPRI to develop this? They realized they were getting too low values and they say drop part of this and we'll do something else?

DR. FORESTER: I think that's the case.

CHAIRMAN APOSTOLAKIS: Okay. Now, if EPRI does this, why don't you say here don't use HCR? I mean, they're not using it themselves.

DR. FORESTER: Well, at this time they would use it. In fact, that paper argued that you should use it first and only if you're getting out on the tail of the TRC where the values could appear to be optimistic, then you would go to CBDT. They don't use that way anymore. I think in the Calculator it's more of a primary method. But HCR/ORE is still a part of that method.

CHAIRMAN APOSTOLAKIS: Well the Calculator doesn't recommend the method. The Calculator includes the --

DR. ELAWAR: It shows that difference by showing those tails.

CHAIRMAN APOSTOLAKIS: Which one are you using now?

Stick to the microphone, please.

DR. ELAWAR: The majority of our members are using the CBDT -- I know of very few people using the HCR.

CHAIRMAN APOSTOLAKIS: Okay. Well, that's very useful information.

DR. ELAWAR: And the information about HCR having low values and curves are shown people looking for a method would already know that in front of them.

CHAIRMAN APOSTOLAKIS: Very good. Thank

you. That's very useful.

DR. FORESTER: Okay. Next slide.

CHAIRMAN APOSTOLAKIS: Yes, I think we discussed this?

DR. FORESTER: Yes. And you saw some examples of the decision trees in the EPRI presentation this morning. There's examples of decision trees that are used.

CHAIRMAN APOSTOLAKIS: Are these -- oh, 27. Still 27? Okay.

DR. LOIS: Shall I go forward?

DR. FORESTER: Yes, go ahead.

And then here's some examples. This sort of describes there's eight different trees, what kind of issues are addressed by the eight different decision trees.

MEMBER KRESS: Are those given equal weight?

DR. FORESTER: Yes, they are. They're treated as independent. So when you come out at the end of a tree, all the values would then be added up.

CHAIRMAN APOSTOLAKIS: And I would -- if I consider this PSFs, would that be wrong?

DR. FORESTER: No, that would be okay. I think.

CHAIRMAN APOSTOLAKIS: Again, what comes to mind is prospective and retrospective analysis. Look at procedural formally. Visibility and salients of instructions raise keeping aids. I mean, is anybody doing prospective analysis going to come in and say my plant is weak with respect to this PSF. I just can't imagine that. This is useful in retrospect--

MR. PERRY: George, can I make a comment here? This is Gareth Parry from NRR.

These things are not PSFs, they're failure modes. The PSFs underlie the evaluation of the probability of these failure modes. And that's why they're additive.

CHAIRMAN APOSTOLAKIS: What do you mean?

MR. PERRY: The different failure -- I'm sorry. No, they probably are the PSFs. But the individual trees are different failure modes of the human failure event.

CHAIRMAN APOSTOLAKIS: Right.

DR. FORESTER: Now the PSFs are the branches on the trees that feed into the evaluation of those. So it's a little misleading to just say you're just adding PSFs like that.

CHAIRMAN APOSTOLAKIS: So these are not

PSFs? So what are the PSFs? I mean, when you say availability of relevant indications?

MR. JULIUS: This is Jeff Julius.

Generally the PSFs are in the parenthesis. The location and accuracy, for example, are the performance shaping factors effecting the failure mode of the availability of the indications.

CHAIRMAN APOSTOLAKIS: So the parenthesis are the PSFs then?

MR. JULIUS: That's right.

MR. PERRY: And the other things is a description of the type of failure mode.

CHAIRMAN APOSTOLAKIS: Yes. But coming back to the issue of prospective versus retrospective, it seems to me that a lot of this stuff, and not just in this method but in many methods, is relevant when you do a retrospective analysis but for a prospective analysis, probably is not something that people will consider.

MR. JULIUS: This is Jeff Julius again.

Well, we have seen this in their practical application of these. For example, the performance shaping factor for place keeping aids, I think there are people in this room who were with me at operator interviews where the trainer said oh yes, we use the place keeping aids. And then we did the discussion in the plant walkthrough and they said well, we do that in training but in actual practice we don't like to mark up the procedures so we don't use them for an actual event. So in that case we put those factors into the HRA update.

The other example is the procedure layout and the procedure wording. There are cases where in the prospective look ahead you find out that a step may be varied and could be better emphasized graphically. And then later that's a suggested change.

CHAIRMAN APOSTOLAKIS: In some cases I can see that, yes. But in many other cases I'm not sure.

MR. JULIUS: But you're right if you're looking at the general emergency operating procedures, there's a lot of times the indications are designed for the actions of EOPs and the procedures are written to emphasize these actions. So, yes, they're not as useful in the prospective case.

CHAIRMAN APOSTOLAKIS: I mean how do you evaluate whether you have a standardized vocabulary or not? I don't know.

DR. ELAWAR: If I may make a small comment, if I may? During my work the availability or leverage of place keeping aids was very important. It factors heavily into the provision of error and as a feedback to procedural writers, they were adding them really quite frequently. Now I see where very rarely I see an action without a place keeping aid for it as the result of feedback they get from us.

CHAIRMAN APOSTOLAKIS: What does place keeping aids mean?

DR. ELAWAR: The operator has, if you do an action, he will sign for it or initial or put the time. You are guaranteed --

CHAIRMAN APOSTOLAKIS: No, I agree that some of these are useful. But I believe it would be better to either have a few comments that some of these are really more useful in retrospective analysis than in perspective or separate them.

MR. PERRY: I'm not sure, George, that these are directly the PSFs that are on the trees. I think some of these are interpretations of them. Because the intent of those trees was to have decision points that were objective that you could actually measure in the terms of a prospective analysis. It's intended for that. So the question on, for example, I don't remember standardized vocabulary being one of the things on the tree. But things like completeness of information would be. And then this would be assessed against the specific scenario in which you're assessing these things. Because the information might be complete for some scenarios and it might not be for others.

CHAIRMAN APOSTOLAKIS: And that's where the confusion matrix would be useful, actually, right? Completeness means can I figure out from the indication of what's going on, right?

MR. PERRY: Right.

DR. FORESTER: Yes. For attention to indications, you know the workload. There's decisions about is it high workload or is low workload.

CHAIRMAN APOSTOLAKIS: No, workload --

DR. FORESTER: You follow right through the tree. Yes, and there is some interpretation here to represent what was in the trees without representing all eight of the decision trees.

CHAIRMAN APOSTOLAKIS: Okay.

DR. FORESTER: But you can certainly measure. And not all of these would always necessarily be important in a scenario. And other times there may be others that would be important that are not included.

So in terms of strengths and weaknesses, again I thought the use of the causal model, you know it simply requires analysts to evaluate potential causes of error. And that's an important thing in my mind.

And there was an effort to look at human characteristics and factors that would influence human performance and use that as a model to help them to identify where things could go wrong.

Using the decision trees are fairly easily to answer the question. Again, you need to develop a very good understanding of what the context is and what's involved in the scenario. But if that is done, then the decision trees can be used effectively, I think.

And also part of the method, even though there was eight specific decision trees, the method itself recommends analysts if there are other issues or other factors they think could be important, they're encouraged to pursue that and develop and take those things into account. So it is flexible in that sense.

In terms of the weaknesses, again there is no guidance. Because it was originally developed to simply address cases where there was plenty of time, it hasn't been tailored, there is no guidance about how you would use it in terms --

CHAIRMAN APOSTOLAKIS: Dr. Kress just brought to my attention the last bullet of the previous slide.

DR. FORESTER: Okay.

CHAIRMAN APOSTOLAKIS: Which is another red flag for a regulatory.

MEMBER KRESS: No, it was the one before that.

CHAIRMAN APOSTOLAKIS: But that is deliberate violations. Is that what he proposed, the violations and then in ATHEANA they're circumventions or something?

DR. LOIS: That's right.

CHAIRMAN APOSTOLAKIS: Oh, EPRI calls them violations?

DR. FORESTER: Yes.

CHAIRMAN APOSTOLAKIS: So what does that mean potential? I mean, you have information about that that these are the shortcuts people take in their normal operations.

DR. FORESTER: Right.

CHAIRMAN APOSTOLAKIS: But do we have any evidence? I mean, I know there is evidence that they do it, but in terms of quantitative impact?

DR. FORESTER: No. I guess if I had my ATHEANA hat on I'd be looking at sort of informal rules, through discussions you might identify places where they might decide to take shortcuts through --

CHAIRMAN APOSTOLAKIS: But then you have to know what to do with that?

DR. FORESTER: Right. Well, you factor it in just like any other kind of factor in terms of how big of an influence, how frequent it would be and so forth.

CHAIRMAN APOSTOLAKIS: Maybe you should change the word "violation." Circumventure.

MEMBER KRESS: Could you answer that with a yes or no and then it kicks out for a thing for you to add?

DR. FORESTER: Well, it gets down to these other kinds of issues. That's sort of a summary of what the whole thing is about. But there is specific questions to get at whether there's a potential for a deliberate violation or not.

MEMBER KRESS: Oh, you answer each one of them yes or no?

DR. FORESTER: Yes.

MEMBER KRESS: And then you add them up?

DR. FORESTER: Yes. It will be yes or no. That's correct.

CHAIRMAN APOSTOLAKIS: There is always a potential. I don't know how you decide.

DR. FORESTER: But is there any evidence that you might think that was going to happen?

CHAIRMAN APOSTOLAKIS: But the potential is there.

I think we said enough about this matter.

DR. FORESTER: Okay.

CHAIRMAN APOSTOLAKIS: Move on.

DR. FORESTER: Now we're up to the Calculator.

MR. KOLACZKOWSKI: You want us to keep going?

CHAIRMAN APOSTOLAKIS: Yes.

MR. KOLACZKOWSKI: Okay.

CHAIRMAN APOSTOLAKIS: Let's see, are we behind? It say evaluation -- oh, it continues after lunch?

MR. KOLACZKOWSKI: Yes, so we're going to continue after lunch, so I mean we could break at any point. But if you want to keep going, that's fine.

CHAIRMAN APOSTOLAKIS: Well, is this a

method, though, the HRA Calculator?

MR. KOLACZKOWSKI: No.

CHAIRMAN APOSTOLAKIS: It's not a method?

MR. JULIUS: It's a software tool, not a method.

CHAIRMAN APOSTOLAKIS: But it's reviewed as part of it?

DR. FORESTER: Yes. Right.

CHAIRMAN APOSTOLAKIS: How long is this?

MR. KOLACZKOWSKI: Well, we still have the Calculator, SPAR-H, ATHEANA --

CHAIRMAN APOSTOLAKIS: Oh, you have a lot. So maybe we should stop now and continue after lunch?

MR. KOLACZKOWSKI: That's fine. That's up to you.

CHAIRMAN APOSTOLAKIS: Okay.

MEMBER KRESS: Yes, let's eat.

CHAIRMAN APOSTOLAKIS: Good idea. Being unanimous, we will recess until 1:30.

(Whereupon, at 12:19 p.m. the Subcommittee meeting adjourned, to resume this same day at 1:29 p.m.

A-F-T-E-R-N-O-O-N S-E-S-S-I-O-N

1:29 p.m.

CHAIRMAN APOSTOLAKIS: Okay. We're back in session. And we continue with the EPRI HRA Calculator. Is it John or Alan? Alan.

MR. KOLACZKOWSKI: Okay. So we're continuing on with some of the method reviews, etcetera.

Again, the next few slides I'm going to spend a lot of time on. You've heard what the Calculator is. And it uses a various sets of models that you can call on.

CHAIRMAN APOSTOLAKIS: What is that exception that you're referring to. One exception you say?

MR. KOLACZKOWSKI: The sigma decision tree. And we'll have a couple of slides on it. But it is something new that was introduced in the Calculator, so to that extent if you will, there was a method that was sort of introduced within the Calculator and not just using THERP or ASEP or whatever.

Strengths and limitations or weaknesses, if you will. And I think we've talked about some of these already in the previous presentation.

Clearly, I think using the Calculator being a software tool, having prescribed windows that you walk through, etcetera, is certainly going to help this idea of consistency. As we try to comment here, it would make it difficult for an analyst to forget to address something because the screen is going to force you to basically say, oh, I got to think about this. I have to decide what I want to do about this PSF or that PSF. So it's going to help in the consistency area. It provides some very traceable hard documentation when you're done, which is obviously good for subsequent reviews as well as going back to whatever you did five years ago and looking what you did and why you made the decisions you made. And that's very good.

There is some flexibility allowed to make changes to some of the basic model and data, although I think they would agree that that's really not encouraged. They really want you to stay pretty much consistent within the data values, etcetera, there. But if you have a good reason to not use, let's say the .03 basic human error probability that's maybe built in the THERP model or built into the ASEP model and you want to use something else, there are some free format fields, if you will, where you can put in or change that value if you have adequate reason. And, hopefully, you would document that reason.

On the weak side, although --

CHAIRMAN APOSTOLAKIS: Let me make a comment on this because I think it's relevant to SPAR-H as well.

MR. KOLACZKOWSKI: Okay.

CHAIRMAN APOSTOLAKIS: I think we have two competing, I don't know, benefits perhaps. On the one hand, of course, standardization is a good thing. At the same time we're trying to standardize something that is so subjective and should be flexible. And the question is where can we find the optimum, okay, so you don't constrain the analysts or the analysts could use judgment depending on the context or whatever. At the same time, of course, you don't want to have an open field where anybody does whatever they please. So it's really a difficult decision, you know.

MR. KOLACZKOWSKI: It is. I think you've summarized HRA almost right there. I mean, that's what it is. Where I think we're looking for standardization, some amounts of constraints and yet not so constrained that when you're dealing with the deviation scenario, as ATHEANA would say, you can move outside the normal and do something different.

CHAIRMAN APOSTOLAKIS: That's true. Yes. Yes.

Okay.

MR. KOLACZKOWSKI: Okay. Weaknesses, we see it although proper training is encouraged. And you've heard a lot about that and whatever. And, again, this isn't so much a problem of the Calculator itself. Again, it's this inherent human nature, we all want to be lazy I think at times, and when you have something that's very easy to walk through it at least is the potential that you can misuse it if you're not properly trained on its use and whatever. And I think they are making attempts to avoid that as much as possible, but clearly --

MEMBER KRESS: The other options make it too hard to use.

MR. KOLACZKOWSKI: Yes. And we know there's a method that people would claim makes it too hard.

CHAIRMAN APOSTOLAKIS: They've done it with the nuclear weapon.

MR. KOLACZKOWSKI: Yes. It's been equated to a nuclear weapon, I believe.

CHAIRMAN APOSTOLAKIS: Not the same people.

MR. KOLACZKOWSKI: That's true.

DR. RAHN: You see you have the horns of a dilemma --

CHAIRMAN APOSTOLAKIS: Microphone.

DR. RAHN: You see you in the horns of a dilemma, you know. If we make it too easy, everybody can use and standardize it that's a weakness, but if we make it too hard that's a weakness, too.

CHAIRMAN APOSTOLAKIS: Absolutely. Absolutely.

DR. RAHN: So finding that middle ground is always a challenge.

DR. ELAWAR: If I may say, at my plant is a person not trained for it, we may well use our accreditation. That's a very important thing for us. So it seems to me that this really should not be a -- because I don't believe people who are not documented as being authorized and knowledgeable in using in doing HRAs, they usually do not use it.

CHAIRMAN APOSTOLAKIS: Well, I think everyone that agrees that the team that's doing these has to include an HRA specialist.

DR. ELAWAR: Yes. Yes. I would never expect somebody --

CHAIRMAN APOSTOLAKIS: Not that we are

trying to create business for these guys, but --we are not. It's important.

So when is a person qualified? Having done it once, twice?

DR. ELAWAR: At my place I can say, I can't speak for industry, we have a lesson plan written that's for you to be an HRA user, a practitioner, you have to go this, this and this and you have to pass a test to make sure that you --

CHAIRMAN APOSTOLAKIS: A test? That's an interesting thing to hear. Okay.

DR. RAHN: If I can expand a little bit. Again, Frank Rahn from EPRI.

As industry progresses, as tools progress, as computer systems progress it's now possible, in fact if you look at a PRA, make it almost automatic in terms of updating. What I mean by that is typically data resides in things like system notebooks, resides in the PRA itself, it resides in procedures. And to the extent that we can, that the technology exists, to do this essentially have hyperlinks between, let's say a procedure and the PRA simply by almost pressing a button and operator checking as we go along.

As an example if we change the procedure where, let's say, a time allowed for a certain action was 35 minutes instead of 30 minutes. If you had the proper hyperlinks so that everyplace that was referenced, not only cross referenced in the procedures but cross referenced to the PRA and the system notebooks, you'd be able to identify what calculations had to be updated. That example you may, since we push a button that says update the HRA Calculator, which then changes the proper point in the HRA Calculator reflect we've gone from 30 minutes to 35 minutes, which then calculates a new basic event, basic event probability, puts that in the PRA and you're finished.

So this is really an important thing we as an industry looking five and ten years out need to grapple with in terms of how do you do that in way that allows for: (a) a living PRA, allows efficiency of the PRA team, if you will, which includes the analyst to do this on a timely basis and yet do this in a way that does not introduce errors and think what particular weakness was addressing, lack of thinking on the part of the analyst as to what it all means in the end.

CHAIRMAN APOSTOLAKIS: Which brings up another point. I mean, we're interrupting your allotted time.

MR. KOLACZKOWSKI: No, that's quite all right.

CHAIRMAN APOSTOLAKIS: A question that the ACRS has been struggling with the last several weeks because we're writing -- you don't know that, Frank, but we're supposed to write a report to the Commission on the research programs of the Agency. And we do this every year. Every other year it's a more detailed report.

One question that was raised is what would we like an NRC staffer to look like? I mean, what capabilities and tools we would like that person to have ten years from now.

So if we focus now on HRA, what would be an ideal practitioner of HRA ten years from now. What do you think that person would be?

MR. KOLACZKOWSKI: Do you want me to take a stab at that?

CHAIRMAN APOSTOLAKIS: Sure.

MR. KOLACZKOWSKI: You know, my background is more I'm a system engineer. And actually the early part of my career was I was designing nuclear power plants and stuff. So I come from a designer --

CHAIRMAN APOSTOLAKIS: Which ones?

MEMBER KRESS: So you're the one to blame?

CHAIRMAN APOSTOLAKIS: You're the one.

MR. KOLACZKOWSKI: Yes, I did some of the control and design on Hope Creek and I don't know what else.

CHAIRMAN APOSTOLAKIS: Okay.

MR. KOLACZKOWSKI: And I'm a recent change over into HRA, maybe in the past, I don't know, two, three, four years, five years. But I'll tell you, one of the things that I felt I needed to learn to become an HRA person, and I'm not sure I've even become one yet, is really understanding some of the underlying behavior science stuff what has been to me very helpful to understand how we go about modeling the human and why we model the human the way we do, etcetera.

And so I think that to use any of these methods correctly, if I can use that term loosely, I think you have to have a basic understanding of behavioral science's approach and so on and so forth, which a typical system engineer or a typical utility person is not going to have. And so you have to train them in some of those underlying sciences, etcetera, that really all this methodology sort of sits on. And I think without that underlying knowledge it's like building a house without having a good foundation.

And that's when you can start misusing these things, etcetera. So that's one thing that I would offer, is that I think if you expect an NRC staff person to review submittals and look at the HRA aspect, I think that person has to have at least some basic understanding of the behavior sciences and so on and so forth and why we break things up into a diagnostic and implementation phase that most methods use. And why we think that's adequate and so on and so forth. I think having some of that basic understanding to me is vital.

So, I've only given you a partial answer, but --

CHAIRMAN APOSTOLAKIS: Yes. Yes. No, I think also what Frank said is very important. I mean, the ability to do these calculations quickly and see the impact is also very important.

But speaking of time, by the way, I'm not sure that there is a model that will tell me -- maybe will tell me, but how believable is it, if the available time goes from 35 minutes to 30, can we figure out now what's happening? And maybe 35 to 30 is not a big deal, but if it goes down from six to four, it is a big deal. And maybe that's one area where we may want to think about.

But I think, yes, these are very good points. And hopefully in ten years we will have fewer models that are acceptable by the community. Not because we declare them acceptable. The community decides that models A, B and C do capture the important elements in most of the situations so people will start using those. I think that is very important, too. Because right now still we have a lot of models. And I think your document here that we're reviewing right now takes a good step toward that. Because, you know, it's a first time that it is in one place, the comparison of models against some criteria that we have reviewed before.

Okay.

MR. KOLACZKOWSKI: I think we'll just move on.

CHAIRMAN APOSTOLAKIS: Yes. Yes, you made some comments. What's the next one?

MR. KOLACZKOWSKI: We do want to make a few comments about the sigma decision tree, which again is a unique aspect of the Calculator that wasn't in the --

DR. LOIS: So we're done with the limitations here?

MR. KOLACZKOWSKI: Huh?

DR. LOIS: You covered the weaknesses?

MR. KOLACZKOWSKI: He can read them.

CHAIRMAN APOSTOLAKIS: Yes, we can read them.

MR. KOLACZKOWSKI: Is there any you want me to discuss? Yes, we're done.

MR. KOLACZKOWSKI: I'll just make a comment on the fourth bullet on the weakness side where the documentation with the Calculator discuss a lot of PSFs but didn't really quantitatively treat them. You're hearing now that in Rev. 3 that's being addressed. So, again, improvements are being made to help to trying to deal with some of this stuff on the weak side.

We did want to make a few comments, though, about the sigma decision tree. And John's going to discuss just the next two slides on that subject.

DR. FORESTER: Yes. Well, this sort of follows the HCR/ORE approach. And this is something that was added to the Calculator to be used to HCR/ORE. And the idea was to have this sigma decision tree so they could address, they could derive some standard deviations that would be able to incorporate some of the plant-specific effects related to training, procedures and things like that. So it was trying to include the ability to address some PSFs. But it follows straight from what was included in the original HCR approach, which the ORE experiments indicated those weren't reasonable to include those in the model. I guess they were nonpredictive was the implication.

So now they're being added back in, it wasn't really clear to us what the basis for adding those parameters back into the monitor.

CHAIRMAN APOSTOLAKIS: So my understanding is that the industry will have a chance to comment on that?

DR. FORESTER: Yes. But we were just concerned that --

DR. LOIS: In a month.

DR. FORESTER: There didn't appear to be a real basis for the standard deviation. There's assumptions that are made that there was no evidence for why to support those assumptions. And, again, we thought those particular parameters had been invalidated in the original ORE studies. So we were just concerned about seeing those added back into the model again.

CHAIRMAN APOSTOLAKIS: That's a sigma,

right?

DR. LOIS: Is that it?

DR. FORESTER: If that's enough, that's just the point we wanted to make.

MR. KOLACZKOWSKI: So now we're going to move on to SPAR-H. We're going to hear more about SPAR-H, so again I'll go through --

CHAIRMAN APOSTOLAKIS: Are these comments you're about to give us come primarily from Jeff?

MR. KOLACZKOWSKI: Again, with the caveat that essentially Jeff provided the initial comments in his review. We had that meeting. We got some more comments. We've reflected those comments into this version, but for instance Jeff has not seen now the latest version.

CHAIRMAN APOSTOLAKIS: So if you think that they distorted your views, please speak up.

MR. KOLACZKOWSKI: Absolutely.

DR. FORESTER: And you may not agree with everything we've said at this point. We've gotten other comments from other people since that time, too.

CHAIRMAN APOSTOLAKIS: Okay.

MR. KOLACZKOWSKI: It's going to sound like a broken record, I guess, but SPAR-H, again, treats error as a diagnostic part and an action part.

Interestingly enough, it does not classify or really distinguish between pre and post-initiator events. You basically go through the same process and even use the same PSFs whether you're analyzing a pre-initiator or a post. So it doesn't really distinguish between the two and, in fact like I said, doesn't even use that classification scheme within its framework.

And just to keep in mind about what SPAR-H was originally set up to do, it was to provide reasonable estimates for regulatory uses, particularly in evaluating the risk of plant events and also as something to be used in phase 3 of the SDP process.

Next slide.

I already mention they look at human failure as a diagnoses contribution and an action contribution. Each is quantified separately. You add it together, you start with a generic rate that gets modified by eight PSFs. It sounds a lot like THERP and some of the other ones that we've talked about, if you will.

Wanted to note on the last bullet here that the error rates and their adjustments to some extent come from review of all the other HRA methods and the values that they provide as sort of a means to ensure some, and I use the terms loosely, validity.

Perhaps it might be better to say consistency with the other methods. So some amount of validity, if you will, has been applied to SPAR-H to say does it give values that I would expect to get similarly using THERP or using ASEP or using some other method?

Next slide.

I think I've already really mentioned these. You start with generic error rates and then you apply the different PSFs. There are some adjustments that you can make. For instance, I just want to call out in the last sub-bullet under the second main bullet, additional adjustment made if there are three or more negative PSFs. This is trying to account for some of interaction that if you're starting to get a number of negative PSFs being applicable, there's some further adjustments that need to be made just so you don't end up with an error rate greater than 1, for instance.

Later on there are further adjustments made for dependencies among tasks. That can be done in the SPAR-H approach. The result is treated as a mean value with an uncertainty.

Next slide.

CHAIRMAN APOSTOLAKIS: It's interesting that the comments here on page 145 it has to do with this pre and post-initiator. It says assuming that the pre-initiator human failure events will be classified as action failures, SPAR-H will assign a nominal HEP of ten to the minus 3. This value was selected based on a review of existing methods. As noted earlier, this is significantly lower than nominal HEPs from ASEP. I guess later on we will be enlightened why that is so? Why they're significantly lower?

MR. KOLACZKOWSKI: Well, and again, that's the first number -- that's the number you start with and then as you apply as the eight PSFs, that number could end up coming up.

CHAIRMAN APOSTOLAKIS: And then there is another criticism.

MR. KOLACZKOWSKI: Well, I don't know if that was a criticism as much as just to say that's a statement of fact, I guess.

CHAIRMAN APOSTOLAKIS: Yes. It's a statement of fact.

MR. KOLACZKOWSKI: They start with that number and then they apply the PSFs.

CHAIRMAN APOSTOLAKIS: SPAR-H reads the PSFs as independent and does not quantitatively consider interactions among PSFs.

MR. KOLACZKOWSKI: Although, again, if we go back to the previous slide. Just like we saw in the Calculator, you analyze each PSF and it becomes a multiplier on basic HEP. So as you multiple these together, they're being treated independently. However, even in SPAR-H when you get to the point of having three or four negative PSFs, there is an adjustment made to, if you will, account for some dependencies among those negative PSFs. So that statement has sort of an exception to it.

And further, when you finally get to looking in terms of dependencies among tasks, again to some extent you're treating interactions, although in this case among two different events. But, yes, if you're just going through the quantification process, the PSFs are treated as independent.

DR. FORESTER: Which is actually important. You know, there can be interactions and the effects if one PSF can change given the presence of a certain levels of another PSF --

CHAIRMAN APOSTOLAKIS: I mean, short available time usually raises the level of stress, does it not?

DR. FORESTER: Right. And actually, you know, they have a discussion of that issue in the document. It's not a real specific treatment of a lot of it.

MR. KOLACZKOWSKI: Bad ergonomics they're going to make the time it takes to do it perhaps harder or perhaps raise complexity. These things are not really independent. I guess what we're telling you is the status of HRA in most methods right now is that we still treat them independently.

CHAIRMAN APOSTOLAKIS: You know, this reminds me of something. Maybe what we can do with these methods, especially the ones that are trying to standardize things, is follow the philosophy of the risk-informed decision making process. Why is it risk-informed? Well, we know that you get the results of the PRA, but then you make a decision using also other things like defense-in-depth considerations and so on.

In decision analysis the current thinking is also that you will get the ranking of the alternative decision options from the formal theory, but you don't do exactly what the theory says. You follow that by a deliberative process where the involved stakeholders evaluate what the result of the formal analysis is and they start departing among themselves whether this is the way to go. In other words, is there anything that maybe has not been modeled, the assumptions perhaps are not a 100 percent valued and so on.

In other words, the trend is to make decisions, regulatory decisions according to the regulatory guide or other decisions using decision theory by ending up to make decisions using judgment, which is informed by the formal analysis. Perhaps here, you know, after we use our standardized methods and so on, we should make an explicit step, include an explicit step that says now you guys sit back, look at what the results of the method are and ask yourselves is this reasonable, does it make sense, do you want to increase the uncertainties for whatever reason. Because as we have all agreed, no method is really perfect. And by making that step explicit, maybe we'll go a long way towards taking away the burden on the analyst of producing results that are really their results. And that probably can ease also the effort to standardize things because you are giving this chance to people to question, to do things, right?

So maybe that's something for the future too, to consider. Because I think in real life this happens a lot, but it's considered an informal step and so on. And what is happening now in other fields is that we are making that step explicit. You will not take the results of a formal analysis and say this is the way I'm going to go. You're going to deliberate on that. And I think the integrated decision making process that's in the regulatory guide is really a good example of that.

So maybe here we can try to do something similar and make sure that at the end the judgment of the people involved, the analysts of course, is really reflected in the distributions or the values whatever it is.

DR. RAHN: There are two old concepts which are just as valid today, I think, as they were 50 years ago. That is first of all the answers from HRA and another analysis are really a guide to your thinking.

CHAIRMAN APOSTOLAKIS: Yes. Yes.

DR. RAHN: It's not necessarily an answer, number one. And number two I think Hans Bayan for a set of documents in '49 that should never use a computer code to calculate anything until you know the answer to one significant figure.

CHAIRMAN APOSTOLAKIS: That's right. That's right. That's exactly right.

DR. RAHN: Both two principles remind you that--

CHAIRMAN APOSTOLAKIS: But we should make those explicit. Because sometimes people, especially people who are not experienced, they might think my God, I used this method, the method says three so it must be three. You know, and it's important to --

DR. COOPER: If I could comment? Susan Cooper with Research.

I think this could also be another part of the answer to your earlier question about what capabilities HRA analysts have ten years from now. And I would add to what Alan said about the base, you know having a firm basis in cognitive and behavior science that they also need to be able to integrate all of the disciplines that play a role in HRA. PRA, engineering, you know thermal hydraulics; a number of different disciplines that actually have input to HRA. And I think more and more of a job of an HRA analyst is not for them to sit back and ponder all of this information and come up with a number on their own, but to be able to integrate inputs and be a facility for debate among people representing those disciplines for them to come to some kind of common understanding and then assign a number as opposed to have one person sitting back and mulling at their desk, you know, what does this all mean.

CHAIRMAN APOSTOLAKIS: No, I absolutely agree with that.

You know, in fact you are familiar. I mean, I think we all have seen that nice diagram that Regulatory Guide 1174 has in the middle integrated decision making process, three inputs and two from the bottom. It would be nice to have a diagram like that for HRA and bring some of these things in the boxes there, maybe one box will ask whether some cognitive aspects have been omitted or whatever else is important. I mean, that will have to be a joint effort with the industry. But I think that would be very helpful, and especially to users. The users will feel much more comfortable, I think, if they knew that yes the guys who are supposed to know are giving me this flexibility to do things.

There is one criticism. This is a criticism, however, in the review.

MEMBER KRESS: Only one, George?

CHAIRMAN APOSTOLAKIS: Not from me. This is from the document.

MEMBER KRESS: Oh, okay.

CHAIRMAN APOSTOLAKIS: That is not up there I don't think. On page 154. There is a discussion of the constrained non-informative prior.

We'll discuss what it's called prior later. But it says here SPAR-H, analysts using SPAR-H should be aware that the C&I prior distribution will in some cases represent less uncertainty than the corresponding log normal distribution from THERP. The C&I prior ignores uncertainty in the mean human error probability produced by SPAR-H, which could be considerable based on analyst-to-analyst.

Maybe it's more appropriate to discuss it with the SPAR-H guys later. But this is an important point. And, again, this point can be accommodated by having this deliberative process again. Because of the analysts and the stakeholders believe that the uncertainty with the C&I is not representative of the state of knowledge, they will have the license to change it and, of course, justify why. I mean, you're not talking about I like it that way. But this is an interesting comment, I think.

That probably comes from your guys?

MR. KOLACZKOWSKI: No. Actually, I think it come from an NRC contractor person, I think.

CHAIRMAN APOSTOLAKIS: Oh, okay.

MR. KOLACZKOWSKI: Subsequent to their initial review.

CHAIRMAN APOSTOLAKIS: Good. Right. It

doesn't matter where it comes from, it's a good comment.

MR. KOLACZKOWSKI: Okay.

CHAIRMAN APOSTOLAKIS: Okay. Great.

MR. KOLACZKOWSKI: I guess we'll move on.

CHAIRMAN APOSTOLAKIS: Yes, we'll move on. But I understand we're going to review ATHEANA now and that's it?

DR. FORESTER: We could address SLIM/FLM, etcetera, if you want. But if you think there's less interest in that, we can -- yes, we could finish up.

CHAIRMAN APOSTOLAKIS: Yes, I was going to suggest that we do that.

MR. KOLACZKOWSKI: Yes, we can do that, George. But just recognize that we also did do a review of SLIM/FLM, etcetera. Because there are a number of utilities that are using that and so we addressed that one as well.

CHAIRMAN APOSTOLAKIS: Although I wouldn't call SLIM a method for human error. It's a method of quantifying judgments, period. All right.

MR. KOLACZKOWSKI: Okay.

CHAIRMAN APOSTOLAKIS: And it's based on another major assumptions.

MR. KOLACZKOWSKI: Okay.

CHAIRMAN APOSTOLAKIS: A curve.

MR. KOLACZKOWSKI: Yes. Okay.

DR. FORESTER: Okay. ATHEANA. And as we've said before, Jeff may not agree with all the conclusions here. So things have been added.

CHAIRMAN APOSTOLAKIS: The arrogance of this. The arrogance of these things.

DR. FORESTER: But it will reflect these initial inputs.

CHAIRMAN APOSTOLAKIS: Everybody knows the article, right? Look at that. No citation. It's from the article that we all read at night before we go to sleep.

DR. FORESTER: It's in the paper.

CHAIRMAN APOSTOLAKIS: It's in the journal, I know.

DR. FORESTER: Yes. No, it's in this paper, too.

CHAIRMAN APOSTOLAKIS: Okay.

DR. FORESTER: Again, we've talked about a lot of what ATHEANA does already. But there is an emphasis in ATHEANA to address in the identification modeling parts of doing an HRA, which goes beyond a lot of just qualification methods. And I think it does it a little bit differently than the way say, SHARP1 treats it and so forth. And it addresses errors of commission. And it does in principle the same concepts can be applied to pre-initiators.

MR. KOLACZKOWSKI: I think the last bullet's worth mentioning.

DR. FORESTER: Okay. Although there has been an emphasis in ATHEANA to identify the error forcing context, I think at some level that's been misinterpreted in terms of how broadly what we want all that to include. The intent is to address both the nominal case and the deviation scenarios. So we want to go beyond just the average type of scenario, the nominal scenario, but we do want to address that also. So we think context and the development of context is important for that case also. It's not just identifying the bad actors that are going to lead to HEPs of 1, but whether the conditions that could also make more the nominal case a little bit harder, or just to be able to understand the nominal case appropriate, the kinds of information you get within an ATHEANA we think is important.

CHAIRMAN APOSTOLAKIS: I think in that respect you're very similar to the EdF method?

DR. FORESTER: Yes, I think that's true.

CHAIRMAN APOSTOLAKIS: They don't go to

context, but it's really the same thing. The same thing; very similar.

DR. FORESTER: Yes.

The next slide. Again, just reiteration that we do try to take a behavioral sciences view, although I don't think it's right to say other methods don't do that also. We did try and focus in on the stage model of information process and consider that different kinds of factors could influence different stages. So that's sort of one of the underlying models of ATHEANA is to try and address that model.

Let's see. In terms of the data, obviously there's no underlying database that we use since we rely on an expert judgment process for quantification.

The data is essentially the information that we gather using the ATHEANA search process and the experience that the analysts bring to the table and their judgments essentially. So the data is collected as part of the process. And ATHEANA in training analysts if you're going to do a PRA at a plant or an HRA, the people that are going to be helping the process we try and provide training for those people on ATHEANA and what some of the important aspects of both behavioral science and industry experience that we think is important. So that's the sort of the data of ATHEANA. There's no numbers explicitly provided in the process.

CHAIRMAN APOSTOLAKIS: Would you remind us what NUREG-1624 is about?

DR. FORESTER: That is the ATHEANA document, the ATHEANA NUREG.

CHAIRMAN APOSTOLAKIS: Oh, okay.

DR. FORESTER: Okay.

CHAIRMAN APOSTOLAKIS: Well, I thought you meant -- isn't there another report where there is an evaluation of human errors of helping observe? There's a fairly detailed -- for shutdown? That was years ago.

DR. COOPER: Yes. That was 1698. That was shutdown. There are actually four NUREGs that have been published.

DR. FORESTER: This describes the ATHEANA quantification process. Again, we use a formal facilitator led expert judgment process. Again, we want to have people, you know operators and trainers, people knowledgeable about how the plant responds to situations, familiar with procedures and understand what will be going on in the scenarios. You know, we have the hands-on kind of information and the other kinds of information we would gather using ATHEANA.

We don't have a preset list of PSFs, although there is guidance in there about the range of factors that do need to be considered.

And there's an emphasis on, again, taking the factors that are addressed, the context that's been identified that seems to be the important drivers, but considering everything together so you have a chance to look potential interactions. And you want to identify the factors that this may normally be something important but in this context this other thing sort of renders that one unimportant. So, again, unless you consider them together in a more holistic way, which is sort of the basis of what we want to do, by doing that you'll develop a better representation of what the important drivers for the scenarios are.

And then in obtaining the HEPs in the quantification process, we do try to develop a distribution for the human error probabilities. So we don't start out with a point estimate. The idea is to try to develop a distribution, considering both aleatory factors and epistemic uncertainty in developing that distribution. So the idea is it's not a generic, your error factors and things like that, are not generic. We try to develop, use the important factors identified by the analysts to help develop that distribution.

CHAIRMAN APOSTOLAKIS: So the price you pay for that it's difficult to use, is that right?

DR. FORESTER: It's perceived as being that way, yes.

Okay.

MR. KOLACZKOWSKI: You'll notice, George, we do have weaknesses on this one.

CHAIRMAN APOSTOLAKIS: Only because Jeff reviewed it.

MR. KOLACZKOWSKI: Okay.

DR. FORESTER: I think Jeff would probably agree it's one of the few that --

CHAIRMAN APOSTOLAKIS: I must say, though, I was really pleasantly surprised when I read the report to see these comments on ATHEANA and SPAR-H. Maybe I had perceived notions that ATHEANA would come out smelling like roses and everybody else would be bad. But this is really a very well balance report. Very well.

DR. FORESTER: Thank you, tried.

MR. KOLACZKOWSKI: We tried to be objective, really.

CHAIRMAN APOSTOLAKIS: Don't over do it, Alan. Don't over do it.

DR. FORESTER: Again, there is emphasis on context. Not many other methods have that type of emphasis. Maybe MERMOS does.

DR. LOIS: Go to the weaknesses.

DR. FORESTER: Yes. I'm trying to decide what I can skip here.

The weaknesses, yes. Just like the other methods, at some level particular since you're using expert judgment process, unless you go to the trouble to really understand what the basis for people's judgments are and you document that clearly, textually, the information is there. It describes what the opinions were, why they were made. Unless you do that, there's no basis for the HEPs. So it does require documentation; that's important. If you don't do that, that is a weakness because you had to way to trace it if you don't.

Obviously, the detailed context development, particularly if you get into searching for deviation scenarios, how the plant conditions might vary that could create problems for the problems, that is going to add extra time to the process. There's no doubt about it.

It can be complicated. We're still trying to, hopefully through some of our experience in doing this, provide more efficient ways of doing that. More shortcuts, I guess.

Let's see. And also, as I said, we see it as still should focus on the nominal case also. And maybe in our attempts to try and make sure people were identifying the deviation scenarios and the kind of context that really could cause problems, we think it's also important that even in the nominal case there's a lot of information that needs to be considered, and it should be gathered. And maybe we haven't done as good a job as possible in convening that information.

Okay. That's it.

CHAIRMAN APOSTOLAKIS: John, let me ask a question.

DR. FORESTER: Yes.

CHAIRMAN APOSTOLAKIS: When you leave this room somebody comes to you and says, you know, I was impressed by your presentation and I have this big PRA. I want you to do the human reliability analysis.

DR. FORESTER: Yes.

CHAIRMAN APOSTOLAKIS: What would you do?

DR. FORESTER: What would I do?

CHAIRMAN APOSTOLAKIS: Yes. You would say?

DR. FORESTER: I would say yes.

CHAIRMAN APOSTOLAKIS: Then what would you do?

DR. FORESTER: For certain.

MR. KOLACZKOWSKI: After you say yes.

CHAIRMAN APOSTOLAKIS: I mean, to help you, would you go straight to ATHEANA? Would you do something else first? Would you use the SHARP framework? Would you follow the guidance in the Good Practices. That's a stupid question; of course you would.

DR. FORESTER: Yes, I would. And I would definitely look at SHARP, SHARP1 in particular. I think there's a lot of good information --

CHAIRMAN APOSTOLAKIS: So you would follow the process and say I will form a team that will have such-and-such a person and so on?

DR. FORESTER: Exactly.

CHAIRMAN APOSTOLAKIS: I'm curious, though. After you do that, would you jump into ATHEANA or do something else first?

DR. FORESTER: No, I think the HRA -- you form the HRA team. But I think one thing we think is very important is that HRA is involved very early in the PRA. So that the HRA team or HRA analysts would be involved in building the models particularly related to the human performance issues and included in those models.

CHAIRMAN APOSTOLAKIS: Okay. So you do that with system engineers, right?

DR. FORESTER: Right. Right.

CHAIRMAN APOSTOLAKIS: Okay. You've done that.

DR. FORESTER: Okay. And at that point you're already in the process of identifying context, I think.

CHAIRMAN APOSTOLAKIS: So you would use ATHEANA?

DR. FORESTER: Yes.

CHAIRMAN APOSTOLAKIS: I thought I read somewhere that you guys are recommending that ATHEANA be used because of its complexity and intensive effort, that you would use it only for cases where the human error is really important, which implies to me there is some sort of screening before that. But you are saying you are not going to do that?

DR. FORESTER: I have seen that written. And I guess if you want to do a full blown PRA and you want detailed answers, then I would use ATHEANA. If you want --

CHAIRMAN APOSTOLAKIS: But why would I want detailed answers for every human error, also the human error in the PRA? I mean, those can be what 200 you said? Two hundred. That's a lot for ATHEANA, isn't it?

DR. FORESTER: Well, even if you use ATHEANA that doesn't mean you can't still do screening.

CHAIRMAN APOSTOLAKIS: Using ATHEANA you screen? There is a screening step in ATHEANA?

DR. FORESTER: Yes. To my mind there is. You begin to build the models, you begin to add the events to the models. You're understanding what the context is. You've done some analysis to the point that you could assign screening values to events, reasonable screening values. And then given those high values if they don't show up as being important, then there's no -- I mean, that's sort of part of the HRA process. Then you don't need to do a detailed analysis of those events.

DR. COOPER: Yes. I guess one of the things that we're discovering with technology transfer with ATHEANA is that people have this viewpoint that if you use ATHEANA, you're using everything. And, in fact, ATHEANA provides lots of different things that you don't have to use every time you do analysis.

You don't have to use the search scheme for identifying human failure events every time. You may start off knowing what the human failure events are that you need to quantify. You don't need to go through that process.

The other thing is the deviation search technique. That's basically PRA. You're trying to identify an accident scenario in its full definition but from the HRA standpoint. You may or may not need to do that.

The principal thing that I think ATHEANA provides that's useful to any HRA right now is a perspective. And that is that context is the first thing that matters and then you find out what performance shaping factors are important in that particular context.

And, in fact, if you try to apply any HRA method to a new technology, let's say we're going to look at NMSS spent fuel pool or we're looking at advanced reactors, you don't have a knowledge base with any HRA method. But you want to try to understand what is going to matter, what's going to be risky. And so you start off and you say under what conditions would a person make a mistake. Why would I care. So you start from that point and then you work backwards.

So it's the perspective that's the most important. And then you figure out what other tools you need to use. You may not need to use everything that ATHEANA provides. I mean, ATHEANA provides a retrospective analysis approach as well. You don't need that when you're doing a prospective analysis.

So part of it we're finding out is that we need to be able to try to package these bits, the various things that ATHEANA can offer, and while it doesn't provide a screening approach right now, that may be something that we can do as well.

CHAIRMAN APOSTOLAKIS: But if you take such a position how can we as an Agency say that when it comes to reactor oversight, which is really what we're doing here, right, and we are running this significance determination process, we're proposing SPAR-H which does not use context. But then, you know, we have researchers at the NRC who say that context is everything and you really have to start with that. Do you see a disconnect there?

DR. COOPER: I think for a while we had

more than one thermal hydraulic code we were using also in the Agency. I mean, we may eventually drop one, we're just not at that point right now.

MR. KOLACZKOWSKI: This is Alan.

Also, George, I guess first of all I'd say, no, it's not that SPAR-H doesn't context. But it may not consider context at the level of detail that ATHEANA would say --

CHAIRMAN APOSTOLAKIS: When you're considering PSFs in essence you're trying to simulate, aren't you? That's part of it.

DR. COOPER: Yes. But ATHEANA sort of turns it around backwards. I mean, in most first generation methods you have a situation described by the PRA and you say okay, so how are the procedures, how is the training and kind of a very general sense. And you were pointing this out earlier on some of the trees that we were discussing in the presentations this morning. Who would ever say they had a deficient procedure? You'd fix it, right?

Now, ATHEANA looks at the other direction. Are there conditions under which the procedure doesn't match? And there are. I mean, the procedures are very good. We've tested them out. They're good for 90, 95 percent of the scenarios that we might encounter, but they're not good for 100 percent. What about that 5 percent? Look at those, how bad is it, what could happen, can you get all the way through an accident sequence? So it turns it around. It's not like my procedures are good, everything ought to be fine. It's when could they be unhelpful.

CHAIRMAN APOSTOLAKIS: Well, coming back to John's point. If I were in his shoes and I said, okay, I'm going to apply SPAR-H first because it's easier to use. And then I will identify as a result of this effort five or ten as opposed to 200 human error possibilities that I really have to understand better. Then I will go to ATHEANA for that. Where would I be wrong? And why would that be inappropriate?

DR. COOPER: You're cut might not be right. You're making an assumption about that SPAR-H is going to get the ordering right to begin with. Or even that your PRA -- and your PRA model is basically designed to try to find equipment vulnerabilities, system vulnerabilities and where the humans come in. With ATHEANA does is try to find where the operator vulnerabilities are, where their gaps in knowledge are and so forth. So I can't say for certain whether it would or not. I don't know.

CHAIRMAN APOSTOLAKIS: Well, I could see

a criticism of that approach being that if you use SPAR-H first and then ATHEANA on what SPAR-H has produced, you may missing other scenarios that may come from a detailed examination of the contents.

DR. COOPER: That's correct.

CHAIRMAN APOSTOLAKIS: On the other hand, do you appreciate that what you just said is pretty strong? I mean, how can this Committee now when people come to us and they said we did a significance determination process using SPAR-H, how can we say it's okay when you tell us that it's probably not okay?

MR. KOLACZKOWSKI: Let me make a comment, George. I think we can't really answer your question yet. The parallel I'd like to draw is you're probably familiar with the ARMEA program back in the '80s.

CHAIRMAN APOSTOLAKIS: Yes. Yes.

MR. KOLACZKOWSKI: And one of the things that it --

CHAIRMAN APOSTOLAKIS: Research money in everybody's pocket, is that what you're saying?

MR. KOLACZKOWSKI: Yes, that's what it was. That's right.

And you recall back then we had a number of PRAs and we were beginning to understand what the

CDF was maybe what was dominating, but people had questions like do you have to model the instrumentation circuits in detail or not, are we missing something. And we didn't know. So the ARMEA in part got created to actually well then let's go do a PRA and really do it in all its glory detail, and I forgot, ARMEA it took 2 or 3 years to do, to find out and answer the question do we have to model this in detail or not.

I think we're in the same thing in HRA. If ATHEANA is opening an door that says, you know, you've got to understand context and could we -- could we be missing the actual risk because we want to believe that feed and bleed, we know what the "average feed and bleed" scenario looks like and we have all kinds of methods to come up with the failure probability of failure to go to feed and bleed, and it's .01 or whatever. But is there a 10 percent chance that the scenario could be different enough that the human error probability would go to one?

Well, if your original value was .01 but there's a ten percent chance that the scenario could evolve in a way that would confuse the operator enough in a way that he would totally fail to go to feed and bleed, you're missing the risk dominant sequence.

We don't know if we're missing it until we try it. And I think ATHEANA, to really understand and answer your question, ATHEANA would have to be applied in a probably, unfortunately, a fairly major program to take a number of HRA events that we might typically see in PRAs and have plant cooperation so we can really develop real plant context in terms of labeling, training, procedures. Not just make it up. And try ATHEANA and see do we get a different answer. And if we do, then shame on us; yes, we're missing the dominant. And if we don't, then you start questioning well then when do we need all this detail.

I don't think we know yet. That's my personal opinion.

DR. COOPER: Well, I think there's another piece to it, and it's not just the number. It's what can understand from the analysis. I mean, all of the discussion that we've had today has also talked about gathering of information, the qualitative analysis until you put a number on the human failure event. And the understanding that you can get from the results really with any of the second generation methods or even the cause-based decision tree at sort of an interim point, gives you might insights as to what's going on. And the insights are more credible.

I mean if you get a cutset in which the human failure event is the so called cause from the THERP table is that they skipped a step in the procedure, probably you're going to go back to the crews at the plant and they're saying why would I skip a step in the procedure, that doesn't make any sense. I mean, I know the procedure by heart. Why would I do that?

Some of the more recent methods that are based on event reviews, operational experience and the advances in cognitive science and behavioral science will give you a different reason as to why that error might occur, which you could take back to the plants and say this is why you might have a problem here, and they can understand. And, in fact, they should because that's where -- those are the experts that are going to be used in the qauntification, the trainers and so forth from operations.

CHAIRMAN APOSTOLAKIS: Didn't you use ATHEANA in some fire scenarios, I understand, the last year or two? Some fire scenarios were analyzed using ATHEANA.

DR. COOPER: The pressurized thermal shock studies used ATHEANA. There were four different studies. I don't think they were published yet.

ATHEANA has been used. I mean, we're going to talk about this a little bit later. I mean, it was the basis for some fire HRA PRA procedure.

CHAIRMAN APOSTOLAKIS: Yes, I thought so.

DR. COOPER: And it's also the basis --

CHAIRMAN APOSTOLAKIS: Let's take the PTS. Could that study be the first half of what Alan is proposing? Would it serve as a first benchmark exercise and maybe have data, look at the same scenarios without looking at the ATHEANA results and see how far SPAR-H can go, and then maybe compare those and start drawing conclusions?

DR. COOPER: You could do that.

CHAIRMAN APOSTOLAKIS: I think you guys could correct me before, I mean, and I keep coming back to that infamous European, at that time it was an European community's exercise. But we have to do something about it. That table will not go away just because it's old. We have to replace it by something that shows that we have progressed.

And I appreciate that doing benchmark exercises in addition to being expensive, requires the collaboration of a lot of people. But we must do something about it. And maybe starting small and taking some scenarios that have already been analyzed with ATHEANA, which is the more expensive method, and then have SPAR-H applied, then we can start making progress. Because there may be a way of coming up with a hierarchy that I mentioned earlier.

DR. COOPER: Yes.

CHAIRMAN APOSTOLAKIS: You know, that this model encompasses everything else but as you know, problems, expenses and so on. But if you do this first and you do that second, then you are going slowly the right way.

But right now I agree with Alan. I don't think we have enough information to decide on this. But, you know, your answers, John's and Susan's, I thought were very interesting.

DR. FORESTER: I certainly agree with your point about benchmarking. We really do need to look at. For one thing we need to see why aren't things consistent. I think it'll be important. But taking the PTS results is a little bit different kind of problem, because we've already identified all the contents. Now once you do that, then it could be argued that another method might produce the same kind of numbers.

CHAIRMAN APOSTOLAKIS: It's not just the numbers. I agree with Susan. It's also the insights.

Are there any pathways that you couldn't have identified with different method and so on. So it's the collection of results. Okay. But of course, the guy who uses SPAR-H on this should not be aware of what you guys produced because even if he wants to be objective, he will be biased.

DR. FORESTER: Sure.

CHAIRMAN APOSTOLAKIS: I think that would be a very good start, and then maybe later we can have a broader exercise, maybe through the participation of the industry trying to compare various methods. Because as we said earlier, the EPRI Calculator, I mean it would be nice to have different things trying to use it on the same problem and then come here and say look at this slide, how great it is.

MR. YEROKUN: We hope to possibly achieve that.

CHAIRMAN APOSTOLAKIS: David, did you wan to say something?

MR. GERTMAN: Yes. This is Dave Gertman with the Idaho National Laboratory.

There is a body of situations upon which SPAR is exercised. Now this is the ASP analysis. And I would suggest that what staff and NRC does is get together the relevant information from the event, including they have access to people at the plant, to the drawings, to the procedures and they routinely will call up for modelers to add insights from Idaho. So you really have a team going through what you believe to be the pertinent information.

I would suggest the way you do an ATHEANA analysis retrospectively and the way you do an ASP analysis is not a difference in whether or not one is detailed and one isn't. I think they have a lot more in common than they do that's dissimilar.

CHAIRMAN APOSTOLAKIS: I'm not so interested in retrospective analysis. I appreciate the lessons we learned, but it's really the prospective that is important to us to make decisions.

MR. GERTMAN: It might be somewhat confounded a bit because what SPAR suggests for a search process, if you go to section 4 within the report, it suggests you use something such as SHARP1 or the ATHEANA ten step process for review of context and important elements. So it's borrowing from there because that was not the intent to develop its own search process to finding out what could go wrong. So you have that. If they both applied that way, it's going to be more similar than dissimilar. But it ought to be interesting to see if the numbers through the convergence of consensus expert judgment and the ones we have with base rates adjusted for PSFs come up in findings within let's say an order of magnitude, which would give you a lot more confidence in which either one you went to.

CHAIRMAN APOSTOLAKIS: I still would like to see it too relatively independent applications to the same problem, just to see what we get out.

DR. GERTMAN:: Well, I think it would be very worthwhile.

DR. FORESTER: Sure.

MR. KOLACZKOWSKI: WE are done.

CHAIRMAN APOSTOLAKIS: You're done. The next steps are obvious?

DR. LOIS: Yes. I guess I'd like to iterate that probably as a result of this evaluation, we should develop an SOP or a regulatory guide or both to characterize the methods and the ability for various applications or regulatory uses.

As you see here, we -- oh, I'm sorry. The third bullet here is, George, we're going this year -- next year we're going to address the ISPRA results.

CHAIRMAN APOSTOLAKIS: Good.

DR. LOIS: And for that we hope that we'll work together with industry to come into some kind of a --

CHAIRMAN APOSTOLAKIS: Give them the paper today so they'll have a year to study it.

DR. LOIS: But we also striving towards developing common frameworks within the domestic and international experts.

CHAIRMAN APOSTOLAKIS: Good. Good.

DR. LOIS: And therefore, all of these next steps encompass, to some extent, your concerns and recommendations. Okay.

CHAIRMAN APOSTOLAKIS: So this confirms again, you know, this time thing. I've noticed that ACRS advice is usually heeded a year or so later after it's given. Which is fine.

DR. LOIS: And mathematician works for maybe 200 years later, right?

DR. COOPER: And Mario is noticing Susan's answer. It's not just nuclear, they also have conventional weapons in ATHEANA.

MEMBER BONACA: That was referring mostly to ATHEANA.

CHAIRMAN APOSTOLAKIS: Great. Thank you.

Are we moving on to the next subject, Erasmia?

DR. LOIS: Yes.

CHAIRMAN APOSTOLAKIS: And the next subject is Susan again with Mike Cheok and David Gertman.

DR. LOIS: It's ATHEANA versus SPAR, right?

CHAIRMAN APOSTOLAKIS: And SPAR, not versus. And SPAR.

Now it says here you need an hour and 15 minutes. Okay. Is that true?

MR. CHEOK: Just for the first two slides.

CHAIRMAN APOSTOLAKIS: Okay. Why don't you move up front.

Okay. Dr, Cooper, tell us how bad ATHEANA is.

DR. COOPER: We're going to talk about ATHEANA and SPAR-H today. We're not going to talk in depth because you've heard presentations on this before. We understand that you're interested in hearing a little bit more about it today. And with that in mind, we'll talk about both of those.

CHAIRMAN APOSTOLAKIS: Let's make sure, though, there is enough time for SPAR-H because --

DR. COOPER: No problem. Yes.

CHAIRMAN APOSTOLAKIS: -- we have some comments.

DR. COOPER: Well, that's up to you.

In particular, the focus of today's discussion is to talk about the uses and objectives of ATHEANA and SPAR-H so you can compare and contrast.

ATHEANA, as we've heard described, is full scope in the sense that it has many different tools, if you will, in its toolbox. It's a second generation method. It includes an error perspective, a knowledge-base, has process steps and quantification approach. Its principal purpose is to support detailed HRA PRA evaluations. There are other uses that are either in progress or have been performed that have not been formally described. And it's best demonstrated when it's used to treat special issues that can be well handled by other HRA methods.

SPAR-H is a simplified method. It has modeling and analysis limitations. It's designed to be used with SPAR PRA models. And it's a general and easy to use method.

That's the overview. I will then talk a little bit --

CHAIRMAN APOSTOLAKIS: What does "consistent" mean?

DR. COOPER: I'm sorry.

CHAIRMAN APOSTOLAKIS: Consistent. You

said consistent.

DR. COOPER: I said consistent?

CHAIRMAN APOSTOLAKIS: The very last line. I think I said simple. Simple to use I think is what I said.

MR. CHEOK: And also consistent. Consistent there means --

CHAIRMAN APOSTOLAKIS: Self-consistent?

MR. CHEOK: Basically they're using the worksheet where we have guides for the users to guide them to use the different PSFs and hopefully they would interpret the same situation, the same scenario consistently based on the guides and the guidance that we give them based on the worksheets.

CHAIRMAN APOSTOLAKIS: Okay.

DR. COOPER: With that very brief overview of the differences between the methods, I'm going to go ahead and talk a little bit more about the--

CHAIRMAN APOSTOLAKIS: Oh, no, let's come back.

DR. COOPER: Okay.

CHAIRMAN APOSTOLAKIS: You are saying yourself best used to treat special issues in HRA. Five minutes ago you didn't say that.

DR. COOPER: Well, no. What I mean by

that is that it's fully exercised in those sorts of situations because you're going to use all pieces that are offered by ATHEANA. You'll use the search scheme to find human failure events, you'll the search scheme for identifying deviation scenarios. You'll use the quantification approach. Whereas, in some cases you may not need to identify human failure events, they may be already defined as part of the issue that you're addressing, or it may be that the issue that you're addressing already defines the scenario. That you don't need to search for scenario or the scenario by definition is a deviation. I mean, in other words there is no real nominal case. It's a challenging situation no matter what way you define it.

CHAIRMAN APOSTOLAKIS: One way to interpret this is that unless you really have a special issue where human error is important, you shouldn't use ATHEANA.

DR. COOPER: No, that's not what I'm saying.

CHAIRMAN APOSTOLAKIS: That's not what you're saying.

DR. COOPER: I'm saying that ATHEANA, the NUREG offers lots of different tools for you to use to do different aspects of HRA. If you want a demonstration of all of those tools, then you go to a really tough HRA problem, and that would be a special issue.

Now, it doesn't mean that you wouldn't want to use ATHEANA a more simple situation. It just simply means that you might not use all of the tools that ATHEANA provides you.

CHAIRMAN APOSTOLAKIS: I hear you, but I mean this agency is approving licensee requests regarding power uprates, all sorts of things, without using ATHEANA. Are they wrong? Are we making a mistake or the other methods may be good enough. Who knows?

DR. COOPER: Well, the other methods are based on an understanding of human behavior that was developed principally in the '70s and '80s. The purpose of all the second generation HRA methods really were to address the limitations of those methods and to try to incorporate a better understanding of human behavior. Now if we haven't decided to or incorporate that kind of understanding into what we're doing yet, that's just the way it is right now. I mean, it's only been in the '90s that people like Jim Reason and Dave Woods, and so forth, have come out with some of the base material for understanding human failures and high risk technologies. And, you know, to take that information and put it into an engineering tool, which is what an HRA method is, has taken a little bit longer. And we're now getting into using it in applications. You know, it's not applied Agency wide, it's just the facts.

CHAIRMAN APOSTOLAKIS: In power uprate decisions, as I said earlier this morning, the issue usually is that the time available to the operator has become short. And, again, as I said this morning if I remember one case, it went down from 32 minutes to 29 minutes. I'm willing to grant that this is not a big deal.

When it goes down from 6 to 4, shouldn't they be doing an ATHEANA analysis then? Because this is critical. Instead of six minutes, now they only have four. Shouldn't they be doing a detailed analysis of the context within which these guys are going to operate instead of dismissing it again and saying "Yes, it's a little worse than the 32 versus 29, but you know the probability doesn't change that much." Well, when will it change? When we have one minute?

DR. LOIS: Can I answer that?

CHAIRMAN APOSTOLAKIS: Of course.

DR. LOIS: Should it done, the human reliability analysis be part of that analysis?

CHAIRMAN APOSTOLAKIS: That's what I'm saying.

DR. LOIS: In my mind, and I don't speak for the Agency, I think no.

CHAIRMAN APOSTOLAKIS: No?

DR. LOIS: Because you should not rely on the operator intervention if you have a two minute difference to --

CHAIRMAN APOSTOLAKIS: I'll take it down below two.

DR. LOIS: These are very short times and this is my personal opinion, to come in and say the operator has two more minutes and therefore can handle this action and therefore my reliability I have a ten to the minus 2 human error probability and I can handle that.

CHAIRMAN APOSTOLAKIS: Well, it happened. I think it was from six to four. It was part of the submittal and dismissed it as, yes, we acknowledge that it may be a little more difficult under 31 minutes to 29, but--but--but it's acceptable.

DR. LOIS: This calls more for guidance--

CHAIRMAN APOSTOLAKIS: Why didn't they

scream bloody murder. Why don't you simply say denied, you do ATHEANA.

DR. LOIS: So that goes for guidance to the staff, and this is an SOP that will tell the staff when to use human reliability; what are we bound, are the conditions for doing.

CHAIRMAN APOSTOLAKIS: I understand. Yes.

DR. LOIS: It's not a matter of what method you use is should you.

CHAIRMAN APOSTOLAKIS: No, I --

DR. LOIS: Accept any human error as a--

CHAIRMAN APOSTOLAKIS: Well, if you accepted the six minutes --

DR. COOPER: Any TRC in that time frame is going to give you a very high number. I mean, you don't need ATHEANA to figure out time is important in that one.

CHAIRMAN APOSTOLAKIS: Yes. But I think it was dismissed in a very cavalier way. And I think part of it is that maybe the reviewers were not aware of all this.

MR. CHEOK: George, I think that's one more thing that we need to consider. When we talk about numbers, we're talking about HEPs here. I guess the bigger picture number is how much does this HEP factor into your final conclusion and your final results. I think that's important. If the HEP factors prominently into your final result, then perhaps it's one place that ATHEANA would be useful. However, if it didn't matter much, then it --

CHAIRMAN APOSTOLAKIS: It mattered, because it was singled out and was discussed. It did matter. I mean, it was not a matter of core melt, but it did matter. It was an important measure.

MEMBER KRESS: Yes.

CHAIRMAN APOSTOLAKIS: So maybe a part of the problem here is communication within the Agency that helps. Making sure everybody understands. Not everybody, the people who should understand better that this tool is available and what it can do.

DR. COOPER: Technology transfer is our principal activity with respect to ATHEANA at this point in time.

Okay. I'm going to talk briefly then about ATHEANA. I think we're going to have ended up having talked about some of this already. But principally want to just remind you because we have briefed you on ATHEANA before, what is ATHEANA, why was it developed, how has it been used, how could it be used and what our future plans for with respect to

ATHEANA.

Again, ATHEANA is not just one thing. It's not just a quantification tool. And I think if the one thing I can do today is this, is to tell you that one of the most important things is the perspective. And this is something I was just mentioning. Second generation methods have a different perspective on human behavior. It's different from the older methods that were based on a viewpoint of, you know, nuclear power plants back in the 1970s when ergonomics issues and procedure format issues were important.

It's not just based on nuclear power plants, though. It's based on advances in psychology for a variety of technologies. But it is an important part that underlies the whole method.

There's also a retrospective analysis approach. Within the prospective analysis approach there's a process for performing HRA, there's a search scheme for identifying human failure events, there's a search scheme for identifying error-forcing context, which really is redoing the PRA from the human perspective in developing an accident sequence involving a human failure event. And then the quantification approach, which as Alan -- well, actually John described is not just quantification but the uncertainty analysis is embedded in that.

Why was ATHEANA developed. One of the principal reasons was to improve the state of art of HRA. It was recognized that there were a number of limitations in the first generation methods. It was recognized way back, you know, these were done and identified and papers written numerous times.

In addition to incorporate the advances and understanding why human errors occur and to more realistically represent errors by looking at operational events and getting lessons learned from those events.

Next slide.

As we've talked already a number of times during this morning discussion, ATHEANA provides lots of new tools, some tools are more sophisticated versions of what has already been used in HRA. In some cases there are brand new tools to do jobs that haven't been done before in HRA. But it does provide a full description of how to perform HRA. It has the systematic search process for identifying human failure events. That's one of the really new things that it does provide. Also the identification of the accidents scenarios, the error-forcing context.

The quantification approach, we've

discussed the flexibility of it. And, you know, the expert elicitation process that we have, it hasn't been described as you describe it, George, or do we have a picture, but it does have the HRA analysis as an integrator role or a facilitator of an expert elicitation process where you h ave people from different disciplines and information that is supposed to be shared among those experts. And then they make the decisions about the judgments, if you will, about the human failure probabilities.

Next slide.

CHAIRMAN APOSTOLAKIS: Formal approach to treating uncertainties new? What do you mean by that?

DR. COOPER: The way it treats uncertainty is different in the sense that the way the uncertainty is incorporated in the quantification approach. As John described, a whole distribution is development in the expert elicitation process as opposed to developing a point estimate and then assigning error factors to it.

CHAIRMAN APOSTOLAKIS: Ah, it's new to the community, to this community?

DR. COOPER: Yes. It's borrowed from other places, but for HRA it's a new approach.

We've talked about the uses some already this morning. The pressurized thermal shock, HRA PRA studies, there were four of them. The Good Practices guidance is developed in part on ATHEANA. We also mentioned the joint NRC EPRI fire HRA PRA methodology. It's also being used for two different MNSS projects, medical uses and also in the spent fuel handling. And there have been some applications outside of the NRC also.

CHAIRMAN APOSTOLAKIS: Wouldn't the context that guys develop, wouldn't that be a very useful input to the efforts to the Agency to understand safety culture? I mean, how can you talk about the safety culture in the abstract? If you produce those deviations and give some idea of the likelihood of these, it seems to me those people would benefit from knowing this unless they are dealing only with a very high level of issues. You know, are you going to have a mock up tomorrow and you know about it and you don't do anything about it. But it seems to me that a lot of the stuff that you're producing, first of all, should be effect by the safety culture of the plant but also you should provide very useful input to the people who are dealing with safety culture.

DR. COOPER: I agree. ATHEANA could use

better input on safety culture in the way we do quantification. And we could provide them some useful guidance as well. We've know that for years.

At present we have not been asked -- HRA has not known -- we have human factors counterparts who are participating in that, but HRA has not been an explicit part of that effort.

CHAIRMAN APOSTOLAKIS: Do you know why you have not been asked or ours is not to ask why?

MR. YEROKUN: I have the human factors and the HRA grouping in Research, so there's a connection there somehow.

I'm Jimy Yerokun.

With safety culture, as you know, I mean it's still in the development phase. For example, the elements to be considered what's safety culture, that's a big deal. We watch it now very closely. I have people involved in the safety culture efforts. There's a definite connection, you know, that HRA implications but how do we -- what is the appropriate connection and how do we get HRA involved is still, you know, some of that is being thought of.

I guess the bottom line is the appropriate time to start getting HRA involved. It's not clear. It's not lost --

CHAIRMAN APOSTOLAKIS: Part of it might be the fact that ATHEANA, as far as I understand it, is not dealing with human errors that may create an initiating event of human attitudes. Because, yes, I can -- maybe it's not 100 percent true, but I mean in the ACRS in two or three letters has urged you to consider normal operations and what can happen do to organizations of deficiencies or whatever that may in fact create initiating events.

Your focus, it seems to me, is really even an initiator, what are the context that that created and how things can go wrong. Is that the main focus?

DR. COOPER: I think that's true. I think I would agree with you that the sequence of events that lead up to an initiator are very closely tied to safety culture.

CHAIRMAN APOSTOLAKIS: Yes.

DR. COOPER: They're closely tied. And as a matter of fact, I would agree with I think it's very tied to your comments this morning about pre-initiator events and whether or not certain branches of the tree that we were looking at this morning with the EPRI Calculator are relevant. You know, the quality or effectiveness of independent verifications and so forth basically catching failures so that they are discovered is going to be very closely tied to safety culture.

The occurrence of the initial failures will have a tie, but I think that can probably be captured with data. But whether or not an organization can correct itself before there's a sequence of events that leads to an initiator I think is going to be very closely tied to organizational factors. And without that piece there isn't much we can do.

CHAIRMAN APOSTOLAKIS: So maybe then there is a natural separation at this time, anyway. Because I think the group that deals with safety culture really worries about things like that as a result of Davis-Besse. I mean that's the reason. And Davis-Besse you didn't have an initiator and then the wrong responses, you almost had an initiator. So maybe that's the reason, that there is a natural separation for the time being of the efforts. But certainly at some point there had to be interaction.

DR. RAHN: I have a question, if you'd like, Mr. Chairman?

Are organizational factors and safety culture synonymous terms?

CHAIRMAN APOSTOLAKIS: No.

DR. RAHN: Are they different?

CHAIRMAN APOSTOLAKIS: No, they are not. Safety management, I guess, includes both.

DR. RAHN: Okay. Then the follow on question is to what extent ATHEANA shed light on what we call organizational?

CHAIRMAN APOSTOLAKIS: All programs or work processes and violations and postponing like what happened in one plant where they postponed some maintenance from Friday to Monday without notifying the appropriate people. On Monday there was something else scheduled. And when both took place, there was a passive -- they lost what? 9,000 gallons of water? Whereas if they had done the work on Friday and the other one on Monday, they never would have created. So somewhere there in the organization miscommunication or something happened. And I would say that's not an safety culture issue. That's an organizational issue, yes.

Safety culture has a lot of problems, as you know, and that's really your approach and the Agency's approach are very different. Because you're talking about regulating something that is not concrete.

So we're all learning, there's no question about it.

Anything else, Susan?

DR. COOPER: Just a quick note about the future plans. As I mentioned before, we're really focused on technology transfer right now. We're working on a user's guide that's in draft form that we've just started. In our review process we'll probably be doing a little more editing before we go for some more internal review.

The purpose of the user's guide is to help HRA practitioners who are familiar with first generation methods, to understand how better to use ATHEANA in applying it in an HRA. So there's some bullets here that sort of outline our approach there.

And then I also mentioned the spinoff products, how else can bits of ATHEANA be used, the perspective and so forth. And then, of course, we'll be looking for other applications.

That's all I have.

CHAIRMAN APOSTOLAKIS: Okay. Thank you.

Any questions for Susan?

The next one is SPAR-H. Maybe we can take a break now, huh? Back at 3:15.

(Whereupon, at 2:53 p.m. a recess until 3:18 p.m.)

CHAIRMAN APOSTOLAKIS: Okay. The next

presentation on SPAR HRA, it's also David Gertman.

DR. GERTMAN:: I'm David Gertman with the Idaho National Laboratory. It's my pleasure to be speaking to the topic of SPAR-H this afternoon.

Next slide, please.

First of all, first of all, is why is SPAR-H? Where do we acquire the performance shaping factors as part of the method? Comparisons that were conducted with HRA methods, including quantification. And in comparison with experiential meeting operating experience data.

Next slide, please.

In 1994 in support of the SEP program, there was a very abbreviated approach to HRA that was used to support that program. There were a couple of rules, such as were actions being conducted inside or outside the control room, were procedures being used, means of this nature and just a few values. And staff came back and requested that Idaho, which was INEEL at that time, develop a richer characterization of human performance and give a finer resolution to the calculation of human error probabilities.

So with that, the SPAR-H as it is today, is really ten years in development. The approach has been a continual iteration back and forth with staff, refinements to definitions, ease of use of the worksheets. We use a worksheet driven approach. And we've gone out, of course, for external peer review and external public comment on the method as well.

One of the main drivers for SPAR-H, and this was a reaction to THERP as opposed to other methods, was that it was felt that it was too difficult to apply, it was confusing, it was time consuming and as George has pointed out in the ISPRA benchmark exercises and others, analysts often using that method would come up with different results, more than an order of magnitude different. Because of that they wanted something that could be applied in a similar, more straightforward approach that hopefully would give more consistent answers.

And by that, there's two types of consistency. One is we force the analyst to always look at the same shaping factors and ask the question whether or not it's mostly a cognitive diagnostic activity that we're looking at or an action based activity which could be just following a step in a procedure that's clearly outlined or in the case of maintenance, performing something that was skill of the craft.

Along the way during the developmental

process we were informed by second generation of International Development Activities. The second generation, the first generation with HRA it's really a somewhat HCR modeling. The diagnoses approach, the diagnoses curves in THERP were pretty simplistic, they're not based upon a large amount of data. I like to think of second generation, the first thing that was important was this notion of a difference between errors of omission and commission. At first we used to just model the omissions, kind of like a nonresponse probability. Then we learned by looking at events as a field that the kind of mistakes people were making, there were two types. One were slips where they had a proper idea but just were improper in their execution. The other one was actually a mistaken sense of where the system was and what actions should be taken. So you had this look at omissions and commissions.

And then context became important as the realization of context by the field and manifest in such methods as ATHEANA and MERMOS and others.

So although we were just trying to get the method a little easier to apply for a number of focus areas that we can discuss, we were also informed along the way by ATHEANA in that process. In fact, back in the beginning of the first couple of years of the ATHEANA effort while Idaho was doing this work, Harold Blackman and I and others sat in on some of the reviews of the ATHEANA back in the early days.

Okay. So I should mention, though, the way we approach context is quite a big different than it is in ATHEANA. We can discuss that.

Next slide, please.

MEMBER BONACA: The question I have and maybe staff can answer, but so the intent is to maintain these two different tools? I mean, ATHEANA and SPAR-H? Using them in parallel?

DR. GERTMAN:: Yes. In parallel. I would liken it to say that in statistics we have parametric and we have nonparametric methods. We're not limited to just one method. Same for NDEE world and other aspects like that. I think it's fine to have different tools to be applied for different situations.

We've heard some that says if you're looking at something where you're looking at cognitive vulnerabilities of the crew where they may be set to fail by procedures, the situation and the behavior of systems which might be unexpected, SPAR-H does not determine that for you. It's a search process from

ATHEANA that would help you identify those situations.

Then as we discussed a little earlier, what you could do is you could take a look at what your quantification within ATHEANA would give you compare and contrast that to SPAR. That really hasn't been done. That would be an interesting benchmark. But you would bring in aspects of ATHEANA in either case.

Part of that is we didn't want to go ahead and try to recreate SHARP or the ATHEANA search process because those seemed to be pretty well developed, put together and have been publicly available.

Next slide, please.

SPAR-H. To be truthful, SPAR-H has always been a snapshot in time, we call it an amalgam of other HRA methods. In the comparisons that we did, we looked at methods such as ASEP and THERP, CREAM, HEART and others. And what we did is we didn't really do a validation. That word's been used, and probably inappropriately. What we did was we calibrated the range of effects of performance shaping factors upon base failure rates from behavioral sciences literature and from these other HRA methods. Again, we wanted for staff a simple, easily to use method where the values generated fell within what was acceptable across what was in use at the time.

Also, we wanted to have the flexibility to be able to conduct the analysis in a relatively short period of time, if need be. It's been used in different ways.

It's been used in the development of the SPAR models, over 70 plant models. It's also been used for ASP event analysis, which can be conducted over a much longer period of time, as well as part of the support for the SDP process.

And again, from those different users we've gotten feedback and we've gone ahead and changed the layout of the forums, sharpen the definitions and added some different features to the approach. And we can over some of those, if you'd like. What's changed since 2003 and what's changed since '99 in that approach.

We believe that we've addressed a good enough set of shaping factors in that we do have caveats for more in depth analysis is warranted, that other methods can be used. But right now we believe we have an 80 percent solution. That the eight performance shaping factors that we have are pretty universal and a lot of situations could be mapped to those.

Certainly the SPAR-H method hasn't really been evaluated in situations where fire and floor and the uncertainties are very great. Because we're not sure if some of the base failure rates we have for those situations and some of the range of influence for shaping factors is really accurate or is too limited.

CHAIRMAN APOSTOLAKIS: David, is SPAR-H intended to be a best estimate analysis or conservative analysis, realistically conservative?

DR. GERTMAN:: I would say it's realistically conservative. We talk about the value being produced as a best estimate in the mean for a base failure rate and it's adjusted for the shaping factors. It's less conservative than some of the ASEP approach. And it considers, we probably have twice the number of shaping factors accounted for in SPAR-H than were accounted for ASEP.

CHAIRMAN APOSTOLAKIS: So I can not really consider it a screening methodology that will lead me to ATHEANA later? I mean, I can screen out a lot of things using your approach which is easy. And then if I end up with ten human errors that we're not too comfortable with, then I can go to ATHEANA. Is that something that would be reasonable to do or am I -- I still have that problem we discussed earlier with Susan, that there may be contextual pathways that you have not identified. But do you think that would be a reasonable thing to do, is say within the 80 percent--

DR. GERTMAN:: Within the 80 percent we're not looking at those. And for most of the scenarios we look at, we're looking at average challenges for bad situations, I think you could probably go ahead and do that. But once you get beyond that, you're still going to want to borrow some of the concepts and ideas from ATHEANA. You're going to ask basic questions: I've got errors, do they lead to unsafe acts? What percentage of the unsafe acts might lead to human failure events? That set of questions that ATHEANA asks is still quite bit -- it should be considered.

I think the other way to use the SPAR-H, you didn't say directly link the insensitivity fashion, too, because of my PSFs I come across with some values just quick approximations. I can look and see what the contribution would be if the shaping factors were much worse. But I think you would be able to do that, use it in a screening fashion with a proper stretch strategy. And for those situations where you say I don't believe the original data really envelopes this, I'm going to have to go ahead and run ATHEANA, I think that's from my perspective, not necessarily the staff's perspective, I think that would be a reasonable approach.

CHAIRMAN APOSTOLAKIS: You say that it has been used extensively by the SDP program. What's the phase 3 SDP --

DR. GERTMAN:: Yes.

CHAIRMAN APOSTOLAKIS: -- where they have to do detailed --

MR. CHEOK: That's correct. And that's the tool that we use right now because of timeliness goals and SPAR-H would be the best tool that they would apply.

CHAIRMAN APOSTOLAKIS: Have you found any instances where the licensee disagree with the human error probabilities you're using and they said, you know, you're way off base, and use our model and we get lower numbers. It's not red, it's yellow.

MR. CHEOK: We get it a lot. And -- and if the HEP is the cause of the disagreement, and I guess this what we have been trying to say, is that the SRA will not perform this HEP calculation in an island. First of all, he would actually converse with the licensee. And then in a lot of cases, he or she would actually contact NNR, Gareth Parry for example, Research Dave Gertman, he and she will get a lot of guidance as to how they would evaluate this HEP and in comparison to what the licensee would have.y

CHAIRMAN APOSTOLAKIS: It would be nice to see examples of this. I don't know when we're going to do this. But maybe walk us through cases where you agreed or the difference was not significance or nobody made a big deal out of it. But also two or three cases where there was serious disagreement. I mean, would that be possible to do sometime in the future?

MR. CHEOK: We can make a copulation for you.

CHAIRMAN APOSTOLAKIS: That would be great.

MR. CHEOK: Okay.

DR. GERTMAN:: Yes, the discussions have been spirited across the phone lines. So, yes, there is room for disagreement and nuances of how you model, although we've tried to sharpen the definitions and that was one of the suggestions from the ACRS in the '03 meetings. We think we've done a better job.

There's still instances where it's not perfectly clear as to which of the PSFs should be manipulated.

CHAIRMAN APOSTOLAKIS: I was telling Erasmia earlier that we have to come up with a schedule of the full Committee to review major products of the HRA problem. And as we know, in February we're reviewing the comparison with the Best Practices.

When do you think the full Committee can review this and maybe there you can incorporate a couple of examples of disagreement? Will March or April be a good time frame or you will not be ready then? Because, as you know, the Committee speaks through its letters. So, you know, this is a major piece of work. I think the Committee should -- first of all, the Committee should be familiar with these methods. And second, you know, maybe they problems or whatever.

When do you think? Mike, is that your purview?

MR. CHEOK: I think we would like to discuss this with, I guess, our managers and with the regions and we'll get back to Eric to set up a schedule.

CHAIRMAN APOSTOLAKIS: But this spring

sounds reasonable? I mean, unless something important comes up?

MR. CHEOK: That's right. This spring sounds reasonable for now.

CHAIRMAN APOSTOLAKIS: Okay. So let's see if we can do that in the March/April time frame without another Subcommittee meeting. We can go straight to the full Committee, which as you know, is an hour and a half. Okay? All right.

DR. GERTMAN:: Next slide.

CHAIRMAN APOSTOLAKIS: You have a comment?

DR. GERTMAN:: Okay. The assumptions of SPAR-H, and then I'll add another couple of these just to energize with some of the discussion earlier today.

First we say for most situations, again, we're an 80 percent solutions; most of the cases, most of the behavior you're going to look a simple modeled human behavior is adequate. And ours is quite simply, there's a sensation perception, an initial part of the model, then a short term memory, a long term memory and then a response. It's basically an information processing model getting the documents mapped to these eight shaping factors that we're derived, again, through interaction with the staff and what was in literature and other methods. That's part of the second bullet, really.

Our model is based on human performance and cognition, not on a specific plant condition. We don't differentiate between pre and post-initiators. We say the neurophysiology stays the same. There's basic failure rates and what changes is the environment, the context and shaping factors around the personnel working. So we believe with the basic human performance model we don't have to make that differentiation. What happens is you look at the difference in -- you know, maybe it's not a procedure, maybe it's a work package. You look at the quality of supervision, you look at aspects of command and control as they fit to that particular situation. So we don't make that distinction.

Again for us, we have a more simplistic approach to context. We define it through the application of the shaping factors.

If your search strategy isn't good, then you're going to miss things. And, gain, it's the application of how you identify the errors. Once they're brought to SPAR-H at attention, the quantification falls out pretty straightforwardly.

Again, we haven't used SPAR-H for extreme events where the uncertainty is great and the data are so thin. Again, it would be interesting to see how SPAR-H would do if we have a couple and part of a benchmarking and sent it to those domains and see what kind of findings we got compared to an ATHEANA approach.

In terms of the HCR which comes up a number of times this morning, I'll give my personal opinion first and then talk about it in terms of SPAR-H. I don't use the older version of HCR for anything.

CHAIRMAN APOSTOLAKIS: Yes, Mike isn't using it either.

DR. GERTMAN:: No. We do include the influence of time, but for us it's a PSF like any other. And we talk about if there's insufficient time to do the task, you fail. There's no miracles. We talk if there is expansive time, then you're afforded an opportunity to recover from an error, for other people to come in to bring other resources to bear. And that assessment is made by the team analysts to go ahead and are reviewing that particular HEP.

CHAIRMAN APOSTOLAKIS: So you can tell us when some probabilities will be when the time goes down to four minutes?

DR. GERTMAN:: Yes. If the task takes 3 minutes and you only have 4 minutes, it doesn't look good. We give you a very punitive rate and we'd rather be a little -- it's the no miracles philosophy on that.

What we do, too, as a result of the 2003 comments, we've set absolute minutes. And now we have relative time. You have to two times the amount of time required to do the task, you have more than ten times the amount of time required to do task; we have those kind of thresholds.

CHAIRMAN APOSTOLAKIS: But there is an interesting point here. It's not really the actual time that's available, it's what the operators think the actual time, the available time is. Has anybody thought about? Because if they think they only have 20 minutes when in fact they have 50, they will act as if they have a time pressure of, you know, 20 minutes. And they may do things that they wouldn't otherwise. I don't know how one handles that.

DR. GERTMAN:: For us it would raise the stress level. Because they would see that their perceived ability to do the task in the time allotted would be stressed for them.

CHAIRMAN APOSTOLAKIS: Right. But they will be less, because they actually have longer.

DR. GERTMAN:: Right.

CHAIRMAN APOSTOLAKIS: You see, the calculation is based on what the thermohydraulic analysis says, not on what the operators think they have.

DR. GERTMAN:: That is true.

CHAIRMAN APOSTOLAKIS: Is that correct?

DR. GERTMAN:: That is true.

CHAIRMAN APOSTOLAKIS: Is that something that there is hope to do something about in the future, maybe in your case or in ATHEANA, or -- this is very hard.

DR. COOPER: To do what specifically?

CHAIRMAN APOSTOLAKIS: Usually we are dealing with the available time as it's given to us by a calculation. But as in real life the operators are not going to run any codes. Now, they are trained, they have an idea but isn't it possible that they might think that they have longer than they actually do or less time then they actually do?

DR. COOPER: Yes.

CHAIRMAN APOSTOLAKIS: So it's really their perception that matters?

DR. COOPER: That's true. And perhaps the folks with the Sandia team that did the PTS can help me remember, but I think we ran into a case like that doing the PTS analysis. You're absolutely right. They're not necessarily familiar with or even thinking about what the available time is with respect to thermal hydraulic code. But they do have sort of an expectation --

CHAIRMAN APOSTOLAKIS: Expectation.

DR. COOPER: -- based on their training.

CHAIRMAN APOSTOLAKIS: Yes.

DR. COOPER: You know, simulator exercises or whatever as to how the scenario may unfold and what that means so far as the pace of their activities. And there certainly could be mismatches between their expectations and the way the scenario actually unfolds. And that can be a problem. You know, not just for implementation but also diagnoses, understanding what's going on and then implementation following.

Alan, did you want to add to that?

MR. KOLACZKOWSKI: Yes, Alan Kolaczkowski.

I was going to say, in a PTS we did enter a few cases. And part of the search process in ATHEANA and one of the things that we did in the PTS work was we knew what the thermal hydraulics about how much time it took, but we would ask questions like are the operators aware of how much time they have? What is their expectations as to how much they have? Do they believe they have a real short time? Do they believe they have a real long time?

Because you're right, what really matters is what the operator thinks he has in terms of how much time.

CHAIRMAN APOSTOLAKIS: And there were discrepancies?

MR. KOLACZKOWSKI: And there were discrepancies.

CHAIRMAN APOSTOLAKIS: Interesting.

DR. ELAWAR: If I may make a comment here?

CHAIRMAN APOSTOLAKIS: Yes.

DR. ELAWAR: The timing is somewhat in proportion to the alarm response procedures and the emergency operating procedures. They are time validated by others. So the operator will go without delay and follow their procedures. And the time will roll on automatically, sort of. Because those are time validated.

For example, I use the map code to validate numerous aspects of some alarm response procedures. And say okay, if they're going to have to do those things, do they have the time for it. I do it separately. I say, yes, they have ample time for it. So the operator does not need to worry if they have time or not.

CHAIRMAN APOSTOLAKIS: But then --

DR. COOPER: Yes, but they're validated for a certain percentage of the scenarios.

CHAIRMAN APOSTOLAKIS: Yes.

DR. COOPER: But not all, not the 100 percent of scenarios. And then when you're talking about something PTS where there are differences in procedural guidance so far as when to make the decision between protecting the core, you know, providing feed water, you know worrying about under cooling versus overcooling. And for some plants that we looked at, the decision point was difficult to decide. When do you change your strategy and when you decide, that change can have a very big impact as to whether or not you get into PTS where the end stage is not core damage, but something else. It's actually a fairly difficult situation for an operator in some cases.

CHAIRMAN APOSTOLAKIS: Okay.

DR. GERTMAN:: Okay. Another issue that came up this morning real briefly was about PSFs and their independence. And we didn't have a slide on this. We acknowledge within the document that the PSFs aren't independent, but then as with most HRA methods, maybe the exception of ATHEANA, we treat them as if they are independent because we use a multiplicitive approach. What we do do is we now since '03 have got a correction factor for the presence of multiple negative PSFs. We try to reduce their influence because we know there's some shared variance there.

Unless we know a little bit more about them, the nature of that correlation is difficult to control for it. One of the things we would hope to get out of HERA in the future as time goes by and the analysis of events is the coincidence of these shaping factors so we'll see the correlation of how these things travel together during events and within LERs and other kind of operating events. And that would give us a basis for determining a correlation and then we would know more of the story about the independence or dependence of these factors.

CHAIRMAN APOSTOLAKIS: Do you have a copy of the report in front of you? Have you got the new copy?

DR. GERTMAN:: The new Reg?

CHAIRMAN APOSTOLAKIS: Yes.

DR. GERTMAN:: Oh. Yes.

CHAIRMAN APOSTOLAKIS: Go to page 14. Table 2-3

DR. GERTMAN:: Yes.

CHAIRMAN APOSTOLAKIS: The caption is "Action PSF Comparison Matrix at Power," right?

DR. GERTMAN:: Yes.

CHAIRMAN APOSTOLAKIS: So the PSFs that you're listing at the available times, stress testers, complexity, experience training, procedures and ergonomics?

MR. CHEOK: No.

DR. GERTMAN:: Three more. Fitness for duty and --

MR. CHEOK: Fitness for duty and work processes.

DR. GERTMAN:: Yes.

CHAIRMAN APOSTOLAKIS: I will repeat the comment I made this morning that you really ought to either have two tables or put an asterisk in some of these and say these are useful in retrospective analysis. Because as I look at it and you have procedures and you say incomplete available but poor, now who on earth from a utility will say our procedure are available but they are poor in a prospective analysis? How can you conclude that they are poor?

In the second column when you give the levels, you have to ask yourself can anyone -- if I'm assess that I'm there in a defensible manner. I can see for the available time, for example, say the time is not available. But that's something that objectively you evaluate it.

Stress, yes, sure, you can say something of complexity.

Experience and training, now I have a problem with that. Could anybody doing an analysis will say, yes, yes, user factor of 3 because our people are not trained well? Come on. Nobody would say that.

In retrospect, though, and your example really refers to augmented inspection teams.

DR. GERTMAN:: Yes.

CHAIRMAN APOSTOLAKIS: They decided or they found that the experience of the operators was low. That makes perfect sense. But in prospective analysis, I think that PSF doesn't belong there.

And for procedures, I would say the same thing. How do you know that they are nominal or incomplete? You don't know that when you do a PRA. When you do an STP, you don't know that.

And then --

DR. GERTMAN:: Often the same it true for HMI, unless you can --

CHAIRMAN APOSTOLAKIS: Yes. Yes.

DR. GERTMAN:: You're aware there's a piece of indication that you would like see in the control room that for some reason is absent.

CHAIRMAN APOSTOLAKIS: Yes. And work processes. Poor, nominal and good. What are you going to do? Go over all of their work processes and have experts and look at them and they declare them poor. And then you have a problem, of course, that if they are poor somebody going to want to fix them, right?

DR. GERTMAN:: Yes.

CHAIRMAN APOSTOLAKIS: So it seems to me that in retrospective analysis these three or four, whatever they are, are useful. In prospective analysis they are not. Maybe you can put an asterisk there and have a big footnote that explains that.

DR. GERTMAN:: I would agree. I had a discussion with some of the analysts in Idaho that were developing plant models and they were saying, you know, a lot of these are just nominal. You know, in terms of developing the model, we never go ahead and say the crew is below average that we've never met, that'd be some distribution of crews --

CHAIRMAN APOSTOLAKIS: I remember that. But it seems to me that this stage is critical.

DR. GERTMAN:: Yes.

CHAIRMAN APOSTOLAKIS: And a footnote explaining that,you know, if you're doing a prospective analysis don't worry about.

Like fitness for duty. I think in the text you say on page 18 in fact, you say for example, an objective measure of fitness for duty may be the time in hours since lack of sleep, which has a variable influence on the performance of different people. How on earth will you know that these guys have not slept well. You don't know that. In retrospect the team says, oh gee those guys were working 12 hours.

So I think an asterisk with a footnote would be very helpful here.

Now, since we are here --

DR. GERTMAN:: Yes, I would agree with that, by the way, because it's not used otherwise and they're all used when you do a retrospective analysis for a cross different scenario.

CHAIRMAN APOSTOLAKIS: In the text, by the way, there is another level for the work processes. It says insufficient level. I don't understand that. The only levels here are poor, nominal and good. Is a role missing or -- something for you to think about.

Now, I have --

DR. RAHN: Well, Mr. Chairman, on your comment about fitness for duty, there are very clear NRC regulations in terms of fitness for duty.

CHAIRMAN APOSTOLAKIS: Yes. So what are you going to do when you do the PRA, you say they comply.

DR. RAHN: Of course.

CHAIRMAN APOSTOLAKIS: Yes. So there is no reason to have different levels. But in retrospect --

DR. RAHN: You might retrospect you might that those are deficiencies.

CHAIRMAN APOSTOLAKIS: That's my point.

Yes, that's another thing regarding experience. It's very interesting. On page 23 -- you didn't know we were going to do this, did you?

DR. GERTMAN:: No.

CHAIRMAN APOSTOLAKIS: You're saying experience training included in this consideration are years of experience of the individual or crew. Now, come on, again, what are you going to say? I'm going to do the PRA and I will -- you know, maybe they mix them. I don't know what they do. It's very hard in a prospective analysis to pass judgment of that.

DR. GERTMAN:: You know what I guess -- if you're in a postulation of a particular sequence or event and it wasn't covered in the T-SAR the way it happened, and you know the crew hasn't been trained to this particular type of event, in that instance you may go ahead and be able to say the training is low because it's simply not covered because it's not required. But 99 percent of the time you're absolutely right, it's not going to fall in a prospective.

CHAIRMAN APOSTOLAKIS: An asterisk with a footnote I think again.

DR. GERTMAN:: Yes.

CHAIRMAN APOSTOLAKIS: And then, of course, there is the big question of where do these multipliers come from. And I think the argument here is that you have your multipliers in the third column and then you have HEART, CREAM, ASEP, THERP. But I don't see a pattern. I'm trying to understand what your logic was. And that's why I asked you earlier did you try to be conservative? If you did, then shouldn't your multipliers be higher than everybody else's with maybe some exceptions when you disagree, or what? I mean, I can see for example time available. You are at a high level. If available time is equal to the time required, you multiple by ten, HEART multiplies by 11, but okay. But then when you go to others --

DR. GERTMAN:: Yes.

CHAIRMAN APOSTOLAKIS: -- you are not always more severe. So I'm wondering what the logic was. How did you decide that the multiplier of .1 or .01 is the appropriate one and not .3?

DR. GERTMAN:: What we don't have here is we looked at the multipliers using HRA and we looked at the range of relative effect from behavioral sciences literature as a group, and that's how far the determination was made.

CHAIRMAN APOSTOLAKIS: But there was no effort to be more conservative than everybody else, was there?

DR. GERTMAN:: No.

CHAIRMAN APOSTOLAKIS: Am I missing it? No. So again, the method doesn't seem to be conservative then, but it might be because everybody else was conservative, but we don't know that. So these--

DR. GERTMAN:: It was more of an attempt to be realistic.

CHAIRMAN APOSTOLAKIS: Well, the Chairman uses realistically conservative, so we use that too. I mean, you don't have to overdo it, otherwise you put ten everywhere. But if you can more a case, if you can revisit these and make a case that, yes, we did try to be more conservative than the other guys, there are some exceptions because we judged that it was not appropriate. I mean that's perfect. Nobody's asking to start using and put number mechanistically there. But they are so important that there has to be some justification.

What else do I have here? I have something.

Okay. Oh, there was one that I saw in the Halden experiments and I don't see it here. Maybe there is a reason. High information load. Why was that considered in the experiments and not by you?

DR. GERTMAN:: A different set of PSFs. There's a number of PSFs that have been researched and our feeling is they can be mapped. I'll take a look at the set and see where that one would find. So, we captured in the definitions.

CHAIRMAN APOSTOLAKIS: Yes, but high information load I don't know where it would belong. That was my first thought, too. It's certainly not available time. Not stress. Is it stress? No.

Complexity? Experience, we brought that. Procedures, ergonomics, fitness for duty; I don't see anyone that would come close to that and encompass it.

Now, from what I saw in the Halden experiments this was not a major factor, although they may correct me in the next hour. But I looked at some and they said, you know, high information load by itself was not important. But if you combine it with something else, it becomes important. So why isn't it part of your PSFs? Maybe it's an omission and you're going to think about and maybe put it back in? Again, you don't have to answer the questions now.

DR. GERTMAN:: No.

CHAIRMAN APOSTOLAKIS: But this is something that struck me as I was reading the documents.

DR. GERTMAN:: Yes, i would agree. It's worthy of thought and we'll get back.

CHAIRMAN APOSTOLAKIS: John?

DR. FORESTER: John Forester, Sandia Labs.

I think some of that is covered under the complexity dimension. There's large number of actions required. There's various aspects --

CHAIRMAN APOSTOLAKIS: But that's not information load. Information load is something else.

I thought it was that, but it's not. And I found the definition someplace, which of course I lost. Maybe the Halden guys can help us with this one.

There is a definition, which unfortunately is not up front.

DR. GERTMAN:: You might want to ask it from the perspective of what does it do to the crew this high information load. If it goes ahead and is a function of multiple instruments and annunciators alarming at the same time --

CHAIRMAN APOSTOLAKIS: Yes. Yes.

DR. GERTMAN:: -- and it's impacting the ability to focus attention on the task, then it seems to fall under stress and stressors for us. But I would agree that there's some additional PSFs, and that's where we would put it, stress and stressors. There's probably another one situation awareness is well researched in the aerospace industry, and we don't have that particular label. So there's probably some PSFs we could look at and say this is how it should be mapped in SPAR-H as opposed to adding a whole new PSF that's clearly linked to a combination of stress and complexity, and then we'd be back in a double counting again.

CHAIRMAN APOSTOLAKIS: See, the

combination is interesting, though. Because in their report on page 8 they say -- you don't have to find it. The operators, however, expressed that the information load failures and especially the alarm sounds were disturbing. It also seemed like the total combination of high time pressure and high information load effected the crew's performance more than only high time pressure. In other words, there was an enhancing effect there.

DR. GERTMAN:: Right.

CHAIRMAN APOSTOLAKIS: And maybe that would be a second generation SPAR, I mean where you look at these results and see whether you have covered it. I'm not saying that you should have already, but you know these are some things that you may want to think about.

Then we have this magic. On page --

DR. GERTMAN:: There's so much magic, though. Which page?

MR. BRAARUD: Maybe I could make a comment? I'm Per Braarud from the Halden Project, and later on we're going to present some more about what you discuss right now. But there is a link between how we define information load and the complexity factor in SPAR-H.

CHAIRMAN APOSTOLAKIS: It's what?

MR. BRAARUD: There is a connection between how we define the information load --

CHAIRMAN APOSTOLAKIS: Yes.

MR. BRAARUD: And the complexity factor.

CHAIRMAN APOSTOLAKIS: But if you considered it significant enough to comment on it in your experiments, I would expect these guys also to say something about it. So that's the comment.

Now we go to page 27.

First of all, at the very top when this is the very top four lines at the end of the previous section it says work processes. Okay. Insufficient information, you see that there?

DR. GERTMAN:: Yes.

CHAIRMAN APOSTOLAKIS: And this is the level that is missing from the table that I mentioned. If I go to the table and look at the work processes, there isn't an entry that says insufficient information, which I think will be most of the time you will have insufficient information. But let's talk about the application of multiple PSFs.

You felt the need to develop a formula on page 27 --

DR. GERTMAN:: Yes.

CHAIRMAN APOSTOLAKIS: -- because if you multiplied the various PSFs and then you apply them to the base rate, you ended up with probabilities greater than one, right? That was the reason. And then you argued that if one uses this formula, the probability is always less than one.

DR. GERTMAN:: I think there were two challenges. One is this is an artifact of the method using those error factors, because you do get a probability greater than one and you keep having to say well everybody knows you truncated one. That was kind of messy.

CHAIRMAN APOSTOLAKIS: Yes.

DR. GERTMAN:: The other thing was the feeling you had raised earlier the notion should you be challenging the results and are they credible. In a number of instances, because we were using negative PSFs, we came out with results that we weren't comfortable with as a team.

CHAIRMAN APOSTOLAKIS: Now, what I would do in that case, I would use a deliberative process. And I would say here if you guys do that and you find that you are at a probability of three, go back and look at it, deliberate it, give some guidance how they do it and then assign a value.

The problem with this is that now you have to defend the formula that you know is difficult to defend. I mean, I don't know why it is. And the other is, of course, that if you don't have a formula, you don't end up with a wrong formula. On page 27 it is wrong.

The plus one at the end should be in the denominator. Otherwise --

DR. GERTMAN:: Yes. Yes.

CHAIRMAN APOSTOLAKIS: -- the NHEP cancels out. Okay. In the examples in the next page it's correctly applied. But I would urge you to not do that. Don't introduce formulas that will put you on the defensive and you will say this and that. I mean, this is an incredible formula. It says PSF minus 1, 400 minus one. I mean, 400? The probability should be wondered. I mean -- so

DR. GERTMAN:: If you go to page E-8 or any of the other appendices, the formula is proper with the 1 in the denominator.

CHAIRMAN APOSTOLAKIS: I know. The next page it's correct, too.

DR. GERTMAN:: Oh, okay.

CHAIRMAN APOSTOLAKIS: Well, obviously it was wrong, otherwise somebody, even a psychologist would have caught it.

DR. GERTMAN:: Yes. We don't know why the number was wrong, but we know how it feels.

CHAIRMAN APOSTOLAKIS: To be wrong?

DR. GERTMAN:: Yes.

CHAIRMAN APOSTOLAKIS: Then the examples that you have on page 28 clearly indicate that these things are useful when you do a retrospective analysis.

DR. GERTMAN:: Yes.

CHAIRMAN APOSTOLAKIS: Because you refer to the augmented inspection teams and so on. So my advice there is drop the formula and find another way, behavioral, judgmental way of handling this situation.

Then I must say this section is not explained very well.

DR. GERTMAN:: I would raise a quick comment. I will address it the way you said, but again in terms of keeping it simple and keeping it repeatable, I know when I pick any three people out of the audience with that formula, given the same PSF level assignment, once we make the correction, I know that number that will be repeated no matter who we bring in. Once I make it consensus expert judgment, I'm not sure.

CHAIRMAN APOSTOLAKIS: Okay.

DR. GERTMAN:: But I agree with your comment.

CHAIRMAN APOSTOLAKIS: It's what we said earlier. The competition between being simple and being reasonably accurate. I mean, I appreciate what you're saying, but at the same time you have to defend it now. And I really don't want to start attacking it. There could be a million other formulas that normalize it and bring it below one, right?

DR. GERTMAN:: Yes. Yes.

CHAIRMAN APOSTOLAKIS: So I don't think -- and we have to acknowledge that a lot of this stuff is subjective. But if your performance shaping factors and the elements, the adjustments factors, they take you clearly above one, I don't see any reason why it shouldn't be one, right. I mean, you have high stress, you don't have enough time, your procedures are lousy. It's one. Why would we hesitate to say that.

And since we're on the subject of the report, I have a couple of other comments. Now, on page XVIII, which is the Executive Summary, you say something that surprised me because you guys, you personally did that analysis that showed that latent errors were important. That's the discussion. XVIII.

DR. GERTMAN:: Yes.

CHAIRMAN APOSTOLAKIS: The second paragraph says "The method does not differentiate between active and latent failures. Identification of modeling of human failure as either active or latent is a decision of the analyst. It is thought that the same PSFs and base failure rates are applicable to either type of error" Now, I don't think you believe that. The latent errors are done by other people, organizational problems so it may contribute to those and so on. So I don't think that you should say that. Maybe all you can say is look, the latent error business is relatively new. We are not handling it.

You don't have to solve everybody's problem here. Okay.

Then you try to say something about work processes and there is a paragraph on the next column. I think you're okay, but I mean I'm not sure that they are used anywhere in this context.

I think I have one more comment.

Page 31.

DR. GERTMAN:: Our friend the C&I?

CHAIRMAN APOSTOLAKIS: Yes. I don't know what my comment was. Where is it? Yes.

And also these laws that you -- Hicks law, Stevens law, Phitts law, are these from cognitive psychology?

DR. GERTMAN:: More from behavioral. Cognitive science and behavioral psychology.

CHAIRMAN APOSTOLAKIS: Yes. And these, you are giving these as models that gave you insights when you developed SPAR-H, is that the idea?

DR. GERTMAN:: Yes. That there was a body--

CHAIRMAN APOSTOLAKIS: You're not really using the logarithm with base 2 to calculate anything? It just give you insights, like you say this law demonstrates that the time required to complete the task is an inverse function of the procedure nor accuracy. That's an insight?

DR. GERTMAN:: Yes.

CHAIRMAN APOSTOLAKIS: That's what you're using. I would put those in an appendix because they are really disrupting the flow of information.

I had some comments on uncertainty, and I don't know where they are.

Tell me what you're comparing on page 43. It was not clear to me. Table 3-1 says base rate, 5th and 95th percentile bounds, and then most of the entries don't have bounds. Do you see the table, the last column?

DR. GERTMAN:: Yes.

CHAIRMAN APOSTOLAKIS: So what are you comparing? Anyway, look at it later.

DR. GERTMAN:: It looks like it's the range there.

CHAIRMAN APOSTOLAKIS: But there is no range. Only one entry has a range.

Regarding the uncertainty now, you're developing a point estimate and then you fit this constrained noninformative prior which gives you the larger uncertainty given that you know only the mean, right? That's what you have to do.

DR. GERTMAN:: Yes.

CHAIRMAN APOSTOLAKIS: But then the criticism we saw earlier is that a C&I may not give you the full uncertainty. If you are close to one, you don't even need to go to C&I.

DR. GERTMAN:: Right.

CHAIRMAN APOSTOLAKIS: But if you are away from one, maybe you want to reconsider. Because if you do that, you are saying I really have no idea what the uncertainty is. I know there is some, and I only have a mean value. So I'll use this distribution that this statistician tells me gives me the largest uncertainty.

I mean, if you were to develop that in a different context, if you developed it in -- and all that, where you know you're going to have data --

DR. GERTMAN:: Yes.

CHAIRMAN APOSTOLAKIS: In that case, the exact form of the prior doesn't really matter that much, or in some aerospace applications all they have is a point value, they declare in the mean value and then they say well the nuks want to see uncertainty, put this constrained thing to show them and pacify them.

I think you do injustice to your work to do that because there is so much insight here. Again, why don't you trust people in a deliberative process to put uncertainties and alert them to the fact that the adjustment factors that you have in the table are not -- they didn't come down from the mountain. I mean, there are uncertainties there.

DR. GERTMAN:: Exactly.

CHAIRMAN APOSTOLAKIS: And give a few examples of how you would do it. I think that would be much better than just saying use this distribution, and then you have a criticism in the other report that says, no, the C&I is not always the most conservative or the widest conservative.

What I'm saying is that in some instances in the effort to make this easy to use, maybe you went a little bit beyond the bounds of reason. We have to admit that this is a subjective thing and you are informing the process using the results of the literature, the experiments, the insights people have and you can push it as far as you can, but not farther. Do you see what I'm saying?

DR. GERTMAN:: I do. I mean, I think it's true we mention -- we don't really deal explicitly with the uncertainty around the PSFs. I don't notice too many methods that do, really, or can't think of them. But --

CHAIRMAN APOSTOLAKIS: No. Even Swain just gave bounds based on his judgment.

DR. GERTMAN:: Sure.

CHAIRMAN APOSTOLAKIS: You know, what else can you do? If you give a few examples where you illustrate how your insights can inform the judgment, I think that's good enough.

I mean, we don't have a problem applying these methods, presumably they have some brains. You know, if you inform them, they will do something reasonable. That's my approach. Because otherwise you have to defend formulas that you know cannot be defended, vigorously anyway. And you have -- anyway, I think you understand where I'm coming from.

DR. GERTMAN:: Yes.

CHAIRMAN APOSTOLAKIS: Other than that, fine. Except for the question why this and not ATHEANA, right? But when the full Committee meets, as I say, you know maybe you can tell us how you will handle some of these comments but also examples, the utility, the disagreements and so on. That would be extremely valuable. Because this model is being used in regulatory arena.

DR. GERTMAN:: Yes.

CHAIRMAN APOSTOLAKIS: I mean it's not just an assessment method that is out there. I mean, our guys are using it. And they are very good, by the way. The region people are very good. So they will catch up very quickly if you tell them, you know, this is a judgment thing. You're not talking to innocence.

DR. GERTMAN:: Okay.

CHAIRMAN APOSTOLAKIS: So, I'm done. Are you done?

DR. GERTMAN:: I believe so. I think the last side is self-explanatory.

CHAIRMAN APOSTOLAKIS: Your last slide

says -- yes. It says stop you've told us already. So, yes.

Gentlemen, shall we proceed to the Norwegian presentation? Do we need to break for five minutes to switch language.

How much time do you need? Who is making the presentation? How much time do you need? You have too many slides. I mean, if you need, say, 2/2½ hours then you can start now and we take a break in between. What do you think?

Why don't we start and maybe spend half an hour or so and then take a break.

So, let's go.

MR. BYE: My name is Andreas Bye and I'm working at the HalDen Reactor project. And my colleague Per Braarud will present this together with me.

Okay. So the outline of the talk is to look at little bit on the role of the data in accuracy, our simulator data. Then we will go through the last experiment in our laboratory, the Halden Human Machine Laboratory. And that is the report you referred to, this Halden Work Report 758. And then a summary after that.

So, the role of data here. And actually, the ultimate goal is a PRA for each plant, of course. For HRA methods, you know, it's used for quantification and a lot of other things.

The role of data, especially from simulators, one thing is to inform the quantification and the use of accuracy methods. And the other is to update, help update actuary methods.

Also we have had another role is to update the repositories or database, and we have had cooperation with Idaho and the NRC on the HERA database.

So three points. One is to inform HRA practitioners in the use of HRA methods. One way to inform this is to look into giving data on occurrence of context. For example, will time pressure occur and then in which situations, in which kinds of scenario is this typically occurring when we're running accident simulations.

Subjective and also objective PSF importance can be help there when there's PSF is present. And we'll look into that later how we really can take a look into that.

And also we have seen that scenarios develop differently based on variability of crews. So that if crews, for example, take certain actions early in the scenario, you will get other context later in the scenario. For example, over time you will get much more time available if you do the right actions early on, for example.

And another important thing is to look into influence of context on human failure or human performance. For example, if you say given high time pressure, what is really effect on the operator and the performance of the operator.

One can look into time pressure limits, for example. When should you use which level of this PSFs? When is there another good time? When is there high time pressure? When is there normal time pressure? Based on the results on looking into whether it effects the performance of the operator or not.

CHAIRMAN APOSTOLAKIS: But you are doing one that's called a PSF at the time or two at the time. I thought the idea behind ATHEANA was that there was a whole context that was important.

MR. BYE: We're doing -- when we're doing collecting or looking into the effect of PSFs, we want to look at one-on-one factor at a time to isolate it in order to be able to say whether this factor or maybe one or two or three factors have influence on performance.

At the same time we characterize the total context of our studies, but we don't manipulate other factors. We manipulate some factors and some factors we only describe how they are there.

CHAIRMAN APOSTOLAKIS: Okay. No, that's reasonable. As long as you have in mind that ultimately it's really the combination that matters.

MR. BYE: Yes. True.

CHAIRMAN APOSTOLAKIS: By the way, is there a better word than "manipulate." I know what he means, but manipulate sounds so bad.

MR. BYE: You use the scenario variance, I think.

CHAIRMAN APOSTOLAKIS: Can someone Google it and find a better word? Manipulation carries with a bad connotation.

MR. BRAARUD: Yes, maybe you that have the English has a better --

CHAIRMAN APOSTOLAKIS: I thought you were collaborating with Idaho.

MR. BYE: Okay. The other thing is informing method development. And here we look into part-validation over PSF weights and thresholds. For example, to look into when there are really an adequate time or, for example, how complexity, what are the effect of the performance and being able to adjust the weights, actually.

Also to look into how many levels for the PSFs. How should you sort of distribute this continuous spectrum of values and levels of the PSFs?

Of course, the same for second generation methods if you don't have specific PSFs or specific levels so you can at least have some information on the influence of performance given certain situations.

Interactions between PSFs can also be studied. Typically one can manipulate two factors at a time and see how they interact actually, together.

So looking into variability and distribution in performance and also there has discussion on validation and benchmark of several methods. I think I'll come back to that when we're looking into next steps there. But it has been mentioned that we have an activity or there plans for doing that. We started to discuss that in the workshop in Brussels last summer. Among the Halden Project members, there has been a discussion on this. And they had an HRA workshop one month ago. And some of these members in the Halden Project want to go into this. So we think of taking one step at a time and at least have an international cooperation to do that. We don't want to embark on that ourselves alone.

Okay. Relevance for second generation methods, for example ATHEANA, quality of the insights and context and crew characteristics as well. I talked about the context in PSFs, but there are also quite some things to learn on the crew characteristics from case studies in the scenarios.

And also quality of the insights on plant conditions and deviations from PRA base case scenarios. As we will see later, there are quite -- some of the scenario variance are quite different from the vanilla PRA scenarios.

Also, the third point. Input to generic database repository for use directly in quantification. I thought I would be talking after Bruce and the Bayesian methods, but I think this will be a topic for tomorrow then.

CHAIRMAN APOSTOLAKIS: Yes.

MR. BYE: Yes. So a possibility to use our results in direct quantification of human failure events. We now believe that you should use our results in combination with HERA methods to sort of generalize our results to each PRA. However, if you want to use this also into respositories, and that's one way of doing this. And they can transfer those further on.

So the results of successes or failures or continuous analogy of that can be put in Bayesian models or other data structures.

Looking into frequency of selected action and then specific scenarios. Because we have quite a lot of scenarios. All in all, in the last study there were seven crews, there were five main scenarios, four variance. So there are five times four times seven; that's 140 scenarios. Actually, that's quite a big database for this.

CHAIRMAN APOSTOLAKIS: You know a question that has been raised by this Committee is how the evidence from Norwegian crews or branch crews operating in Norway, how is that evidence relevant to American crews in Texas?

MR. BYE: Yes.

CHAIRMAN APOSTOLAKIS: Do you have any Texans in your teams?

MR. BYE: Not yet. There's three points to answer that thing.

One is that the way we do the studies with controlled variance or manipulations of certain factors where we keep all other factors constant. This is a typical sort of a classical psychological experiment. In that case, we mean that we can isolate the factors that are varied so that if there are differences, systematic differences in the outcome of the human performance, we can say that then the result of the unit performance or the differences in the results are due to the manipulated factors because all we do within subject of science, we will go deep into this later. But all crews run all scenarios so that you know they all have the same sort of computerized setup in our lab. And we know that can say something about if you manipulate such a factor or two factors at the same time, we know that this case the performance difference.

CHAIRMAN APOSTOLAKIS: Can you give us some idea of what kinds of crews you are using?

MR. BYE: Yes. We will go quite deeply through this methodology later, so maybe we could -- but they're licensed operators, I can say that. I think we should go through many aspects of these methodologies later.

CHAIRMAN APOSTOLAKIS: All right. All right. We can wait.

MR. BYE: But that's the first point.

In addition, we also try to dig out crew characteristics here based on case studies of the scenarios. And then you can argue, well we need to have similar operating, for example, culture among the crews to which the ones we want to generalize to.

So the second point is that -- or the operational culture is rather similar between different plants around the world. If you look at plants within one country, there might be as big differences in culture as between countries. We run now, for example, on the PWR simulator. We have Westinghouse EOPs, that's also used in Korea, for example, or all around the world.

Of course, I know that you won't believe that statement. So we also want to get U.S. operators to Halden in order to run scenarios and run studies on our Westinghouse simulator.

CHAIRMAN APOSTOLAKIS: Have you run any experiments with American operators?

MR. BYE: We have not yet. We are working on getting American operators. And --

CHAIRMAN APOSTOLAKIS: When you say "American operators," you don't mean American American. I mean, from one plant.

MR. BYE: Yes.

CHAIRMAN APOSTOLAKIS: People who are working together?

MR. BYE: Yes. Sure. And we need to have people from the plant they're simulating, of course. Because that's really important to have -- and I've been talking to Jeff also in June this summer when we were Washington and talking to EPRI. That might be a connection there to get contacts with the plants.

MEMBER BONACA: So what you're comparing, however, is crews from different countries but following the same procedural framework and process?

MR. BYE: Yes.

MEMBER BONACA: Okay. So the same formality that is used. Okay.

MR. BYE: Yes. We have done quite a lot of studies. And we have a computerized setup in our control room, which is not the one they have in the plants the operators are coming from. Then they have onlog panels and so on.

We have seen that if you talk about differences, functional differences in how the simulator is behaving is more important than actually interface differences on the surface. That might create longer times for reactions and so on, but it does not really create a big confusion among the operators. What is really important is that their behavior and the process is behaving as they are accustomed to back home when they're operating the plant. So it's important to have operators from even the plant we are stimulating or the sister plants or whatever.

I mentioned HERA, that we have an activity with NRC to populate HERA with simulator data. And it can also increase the use of HERA maybe on simulator accident situations. Similar for NARA, actually. They are using data, have been using data from all kinds of studies, also earlier Halden studies and taking this into account.

CHAIRMAN APOSTOLAKIS: What is NARA?

MR. BYE: NARA is the successor of HEART. HEART is used very much in the UK. Developed by Jerry Williams at one point. NARA, is Barry Curvin who is heading the development of that.

CHAIRMAN APOSTOLAKIS: So they are really not nuclear?

MR. BYE: What?

CHAIRMAN APOSTOLAKIS: They are not nuclear?

MR. BYE: Oh, yes.

CHAIRMAN APOSTOLAKIS: Barry is airline--

MR. BYE: He is in your control, but he is contracted by British Energy to develop NARA for nuclear. Yes.

Okay. So what we dive into day is this report, this task complexity experiment. And to get a feeling which PSFs we are looking at. These are the PSFs from the Good Practices. There's ten of them. And as you were into, they're all different definitions of PSFs or context in every method. What we try to do is to explain very clearly how we have defined it, maybe some hints to how that maps into other methods, but not always. That would be the reader to decide that. But the ones we are actually touching upon here is at least time available and time required to complete that including the impact of concurrent and competing activities. It gives information on that.

The complexity of the required diagnosis, also information on that.

Workload and more sort of felt time pressure.

And also based on the case studies we have done of some of the runs here, we can something about crew characteristics.

And also consideration of this realistic accident sequence diversion. I think it gives some information on. So that's up to you to judge when we'll dive into this now.

MR. BRAARUD: My name is Per Braarud, and I work also within the Halden Project.

My background is mainly in psychology. I have been working nearly ten years with simulator studies in our laboratory planning and conducting analysis such studies.

CHAIRMAN APOSTOLAKIS: Are you a psychologist, Andreas?

MR. BYE: No. I'm the only one in the group that's not, actually.

CHAIRMAN APOSTOLAKIS: And what are you?

MR. BYE: I'm an engineer, control theory.

CHAIRMAN APOSTOLAKIS: Okay.

MR. BRAARUD: Okay. Present an example. One part of a study we performed and completed last year. And I will also focus quite a bit on the background for the study and especially the methodology for the study.

And Andreas has already presented quite a lot of background for why we're doing this. I will not repeat that.

So we have selected three factors that we wanted to study how they effect human performance. And these factors, they come from previous work where we have asked operators after completing accident scenarios to rate a set of factors, how would they describe these scenarios. It was, for example, things like if there were many alarms in the scenario, many tasks, did it have time pressure and the need to act on the process and so on.

And by analyzing these data we found that three broad factors can explain the set of factors as a factor analysis.

So these factors we think they describe three important elements that the operator experience during scenarios. So these factors can distinguish different scenarios.

So it's defined such a way that time pressure has to do with how the operator feel. If he feel the need to act on the process, and of course the time available is one element in this definition. And also information load was defined as how much is it to do in the scenario, is there many information elements that need to be taken into account and are there many tasks that need to be operated simultaneously.

We have a third one called masking, maybe that is not even a very good English word, actually. We think about ambiguity about the process situation. Is it difficult, let's say, match the current picture with some idea what is the situation or is it difficult to observe what is the cause for the process symptoms.

And these factors they are not completely independent. If there are much information load, this will also effect typically to some extent the time pressure or the time available.

MEMBER BONACA: I have two questions.

MR. BRAARUD: Yes.

MEMBER BONACA: This study then is only for control room operators?

MR. BRAARUD: Yes, this study is for control room operators.

MEMBER BONACA: The second. Is it focused only on individual performance or also crew performance.

MR. BRAARUD: It is focused on the crew performance.

MEMBER BONACA: On crew performance. Okay.

MR. BRAARUD: I will explain some more. Yes, it's control room and crew performance. Okay.

So the research questions, they were at a general level. How does these factors effect human performance, and we did a methodological choice of how to study this. And this was that we developed a scenario with a main task of interest. By adding tasks to this scenario, we planned to create time pressure, information load or ambiguity or masking for the crew. And the reason for this was to be able to separate the effect of the context being these three factors on a given main task.

And this implies some assumptions. That is, for example, if this additional task will create the effects that we're expecting them to do.

So based on this three factors that give a picture of how the operators experience the scenario, we tried to develop additional tasks that will create this concept or this phenomena.

Okay. This is actually a little bit in the same line. We expected that this additional task, they were designed to create three phenomena similar to those three factors that we previously had identified. So then some more about the methodology for this experiment.

The participants for this study was seven crews and they have three licensed operators. They are licensed to operate the plant we simulate or assist the plant for the for this plant.

CHAIRMAN APOSTOLAKIS: Can you tell us

what the nationality was?

MR. BRAARUD: Yes. They were Swedish.

CHAIRMAN APOSTOLAKIS: All seven? Seven crews?

MR. BRAARUD: All seven crews are Swedish. That is because we simulate a Swedish boiling water plant.

MEMBER BONACA: In Sweden do operator use the same approach to -- do they have symptom oriented procedures, do they follow them literally or is it different? I'm just curious. I mean, you are familiar with the procedure in the U.S.?

MR. BRAARUD: Not in detail. But I will say some about the procedures they used for this study later.

So the configuration of three operators, supervisor, reactor operator, turbine operator. This is the normal configuration for the plant for the control room. In addition, they have two field operators as a normal configuration.

And as I said, they came from the simulated plant or from the sister plant of the simulator.

So just to give a short description, if you look at the mean age, also the distribution for the operators, we can see that this resembles, let's say, an industry with experienced people operating the plant. The two supervisors, they have a mean age of nearly 50 years. Nearly ten years mean experience as a shift supervisor.

Reactor operator mean age of 44 years. Seven and a half years experience as reactor operators.

Turbine operators, 37 years.

So they were quite experienced people.

So this is also a comment to a previous comment that if you compare this kind of data to data previously used for HRA, for example when you base it on psychological experience with, for example, students in let's say simple lab settings, this study is much more close to the actual operation that we want to explain.

So the simulator we used in this experiment, it is a boiling water reactor and it simulates a Swedish boiling water reactor. And this is a quite late generation ABB plant.

The simulator is a full-scale simulator. It's very comparable to a training simulator. And it has a computerized human-machine interface.

MEMBER BONACA: Is it a faithful

reproduction of the control room or it's just -- it's more of a simulator -- or not?

MR. BRAARUD: If you look at this picture, it give you a picture of the control setting in the lab. And this is, the layout is not comparable to the actual plant.

MEMBER BONACA: Okay.

MR. BRAARUD: But the interface was designed to resemble a typical interface for the actual plant. So it's designed based on, for example, their P&ID programs. Their documentation is used as the basis for using the performance, process performance.

MEMBER BONACA: And you have the reactors to the left and the turbine to the right?

MR. BRAARUD: Yes. This shows the reactor operator to the left, the work station. Turbine operator to the right. Supervisor --

MEMBER BONACA: Right here.

MR. BRAARUD: -- closest. And we also have a large screen that present information that should be similar to the overview information that they have available at their plants.

DR. RAHN: Excuse me. Question. Does your Westinghouse simulator also, is that a faithful reproduction of, let's say, a Beaver Valley plant I believe it is?

MR. BRAARUD: Excuse me. Are you asking about the interface?

DR. RAHN: No, I was asking about your Westinghouse simulator. You were talking previously about perhaps having U.S. crews at Halden. And I was wondering whether or not your Westinghouse simulator is a faithful reproduction of a U.S. plant.

MR. BRAARUD: Yes. That simulates a French PWR.

DR. RAHN: Thank you.

MR. BRAARUD: Yes, which is a Westinghouse design from the '70s. The plant is actually quite comparable to at least a couple of U.S. plants. And also the interface is computerized and designed on the following similar principles to resemble how the crew work in a conventional or the actual control room.

Okay. Also something about the procedures. They are actually the procedures for this simulator is copy of the simulated plant procedures. So they are the procedures that the operators are used to use.

There is one difference, and that is that the sister plants, emergency operating procedures are a bit different. And that they use their emergency operating procedure when they run the plant. And this is a procedure set where they have typically normal operation procedures. They have procedures to bring the plant to different stage, typically shutdown/startup procedures. And the procedures for accidents or anticipated accidents, they are evidence based.

MEMBER BONACA: They're not symptom based?

MR. BRAARUD: No. But the emergency operating procedures, they are symptom based or function based. The simulator and the sister plant. That's the package.

Also in addition, they have a special procedure that they call a first check procedure that they run after an event is initiated or if they like to run this procedure to get a overview of other plants.

Also the question of how realistically can a crew run the simulator in the lab. And all experiments they include a training session with the aim of getting the operators knowledgeable and use to using the interface in the laboratory, which is computerized. So there is going through the details of the interface, putting weight on some special features. They also get some information about the scope on the simulation, being that some systems are maybe 90 percent simulated, some of them 95 and such on.

And also the documentation in the control room is aimed to be as similar as what they have at their plant.

And typically we run several training scenarios and test scenarios to see that they can handle the interface in a good way and actually concentrate on the process problems.

CHAIRMAN APOSTOLAKIS: So how long do these crews have to stay in Halden?

MR. BRAARUD: This depends on the different experiments, but in this case they stayed for one week.

CHAIRMAN APOSTOLAKIS: One week?

MR. BRAARUD: Yes. Each crew stay one week.

CHAIRMAN APOSTOLAKIS: Including the training and all that, one week?

MR. BRAARUD: Yes. They use approximately 1½ day to train on the simulator.

We also give them information before they came to Halden. For example, pictures of the process formats so they can be familiar with the interface before.

MEMBER BONACA: But it seems to me that with 1½ day training that you put them under time stress that may effect -- I mean, the lack of familiarity with the system may be of more influence in the lab.

MR. BRAARUD: Actually, we observed that they remarkably fast learn to operate the process through this computer performance.

MEMBER BONACA: Okay. So you feel comfortable that they have learned enough that they are pretty much able to move automatically from one display to another?

MR. BRAARUD: Yes. We feel they are quite comfortable running the plant. There may only be some special issues that if they don't -- let's say, can navigate as good as they should. But that is only rare exceptions. So that's maybe also quite interesting results for computerized interfaces. They learn this very fast.

Also there is a -- of this simulation. We tried to run the scenarios in a, let's say, planned way so that the run is as similar as possible for all the different crews that participate.

And so we have some procedures for the

staff running the experiment to ensure that, for example, failures in the scenarios are set at the same time and so that the starting point are similar for all of the crews.

And also typically there is functions performed during the simulations. One important one is the one on the last bullets, and we are, actually, you can say role playing several of the important external communications that the control room want to make. For example, the field operators are simulating by a person. The control room, they call, use the telephone as normally and say that I want to have a field operator going to that system doing that operation. And this person tried to simulate by himself the time he will think this will take and report back. And operate in the simulator to a work station.

Also the crew can call, for example, on the safety engineer. That's mostly to have the supervisor during the actions he would normally would do in such a situation. They can also call plant management and other persons. But it's actually the field operator is played, it's like most realistically. That's a person doing important actions for the crew.

And also we have some observations of the behavior. Typically we are giving expert -- giving comments during the scenario. And we also record all the room, we do. We record all the communications in the control room.

So this is a setup of the simulation.

CHAIRMAN APOSTOLAKIS: Did the Sweds pay for this? Who paid for this exercise?

MR. BRAARUD: That is the Halden Project.

MR. BYE: This is part of the main research program in Halden that this -- so it's -- there are 80 nations paying for this including the NRC.

CHAIRMAN APOSTOLAKIS: If we are paying, you shouldn't spell behavior that way.

MR. BRAARUD: Maybe it's the UK over spelling it. We have to give them something. You get the results and they get the spellings.

But this actually describes mostly all the method, the background for Halden studies are performed.

Now I will go some more into an example that I performed.

So this experiment investigated actually you could say three elements. The most important one or one of them was time pressure and information load, as mentioned, masking aspect and also one element was it states accident operation further down the event sequence. It's actually an scenario where a previous function has failed for technical reasons and the crew has to get a second function working. It's actually the low pressure coolant injection where the high pressure coolant injection have failed before.

MEMBER BONACA: So the masking is a leakage from the shutdown cooling system?

MR. BRAARUD: Yes.

MEMBER BONACA: Okay.

MR. BRAARUD: Yes. To the left is what we have investigated, and this is implemented in scenarios shown in the column to the right. So I will actually first take the masking as the example. And this was implemented in the scenario that we call leakage from the shutdown cooling system.

So the design of the study is, I mentioned briefly also previously, is that we can call it a base or a nominal scenario where we tried to add tasks to create the phenomena that we want to study. So this context is studied by the scenario variance, wholly different from this base case. Typically called experimental conditions simulator or manipulation, if you like.

Okay. And also it is what we call a within subject design and its such that all crews, they run all this variance of the scenario.

MEMBER BONACA: These crews are coming from a plant?

MR. BRAARUD: Yes.

MEMBER BONACA: So you're not mixing individual from different crews right now. You're taking an experienced crew and put them in the simulator?

MR. BRAARUD: Yes.

MEMBER BONACA: So they know each other?

MR. BRAARUD: Yes. All the members from in a crew are from the same plant.

MEMBER BONACA: So they know each, they're used to work together?

MR. BRAARUD: Yes. Either they are a crew that have worked together at the plant.

MEMBER BONACA: Yes.

MR. BRAARUD: But not always. Sometimes it's what we call a mixed crew --

MEMBER BONACA: Okay.

MR. BRAARUD: -- that come from the same plant but not work together normally.

MEMBER BONACA: All right.

MR. BRAARUD: So they're involved.

CHAIRMAN APOSTOLAKIS: So they were willing to send 21 people for a week, or they didn't stay? They stayed the full week, right?

MR. BRAARUD: Yes. It's actually since there are two plants, there are four crews from one plant and three crews for one plant.

MR. BYE: This is part of the cooperation agreement we have with Swedish participants of the Halden Project. And the main signatory member in Sweden is -- but also the utilities have interest in this. And as part of this agreement, they send some crews. But also it is important to state that both the cres and the utilities see their own interest in this. They are interested in this because they see that it's like additional training for them in a lot of scenarios that they want to do otherwise.

CHAIRMAN APOSTOLAKIS: Frank, do you think that there's a chance that an American utility would send so many people, or you can find sister plants maybe?

DR. RAHN: I am not the one to make that decision.

CHAIRMAN APOSTOLAKIS: I understand. But

do you --

DR. RAHN: It's up to the individual utilities. I think there may be of some interest there. I invite our friends at Halden attend our next HRA meeting, which is coming up in a few weeks.

CHAIRMAN APOSTOLAKIS: This is very interesting stuff.

DR. RAHN: Yes.

CHAIRMAN APOSTOLAKIS: Very interesting.

DR. RAHN: And I think it would be --

DR. ELAWAR: I think there is a compelling reason that they will send, just like that, I don't believe my plant will send unless they find some compelling reason for it.

CHAIRMAN APOSTOLAKIS: And what would the compelling reason be?

DR. ELAWAR: Like for example, suppose an extensive task that will cost them hundreds of thousands of dollars, for example, or people not to pass their NRC tests.

CHAIRMAN APOSTOLAKIS: You get to Norway. That's cheap.

DR. ELAWAR: If I may ask a question? Do the operators have a chance to talk to each other at the end of the day to see what did you do today, and maybe they will ask me that tomorrow? Is that part of the deal? I'm just asking, was that consideration or not.

MR. BRAARUD: You mean if the crew can talk together or --

DR. ELAWAR: At the end of a day were they instructed not to disclose information to each other?

MR. BRAARUD: We ask them, for example, not to discuss the scenarios with their colleagues.

DR. ELAWAR: You did?

MR. BRAARUD: At the plant, for example. So that the crews coming for the next -- the next crew coming next week, should not have discussed it with their colleagues. And we think they respect that.

CHAIRMAN APOSTOLAKIS: They were not all there at the same time?

MR. BRAARUD: No. They are there for a week in a sequence.

CHAIRMAN APOSTOLAKIS: No. I mean, these are all seven crews.

MR. BRAARUD: No, no. That's true.

CHAIRMAN APOSTOLAKIS: So when the crews are finished, they're not supposed to talk to the crew that was going next week.

MR. BRAARUD: Yes. Yes. That's the case.

I think one important reason for the plants sending crews to Halden is that several of the plants in Sweden, they are modernizing, upgrading their plants and this imply that they are upgrading their control rooms. They will get very good experience by running the plant by the computerized interface. So they see this value.

MEMBER BONACA: Sure. Factored into a control design. Sure.

MR. BRAARUD: Yes.

MR. BYE: They get a lot of ideas through this, actually and they say they can use it.

There is another thing also. They are doing -- the operators are doing this on a voluntary basis. And I think some of them do it on their sort of the free weeks when they have sort of daytime service and not have -- and they get paid to do this and so on. And so it would be a week of interesting work in Norway.

MEMBER BONACA: And the operators are not concerned about the feedback that their company may get about their performance?

MR. BRAARUD: This is also an important point. We say that we will not give any detailed feedback to the plant about individual crew's performance or individual operators.

MEMBER BONACA: Because I know the operators are very defensive about that.

MR. BRAARUD: Yes.

MEMBER BONACA: Particularly if you have scenarios that are not completely within their training?

MR. BRAARUD: Yes.

MEMBER BONACA: Okay.

MR. BRAARUD: We run some difficult scenarios, and that is very important that it not be possible to identify the different crews.

MEMBER BONACA: Right.

MR. BYE: They have asked for that, actually, but it's not -- there is another talking about cooperating with U.S. plants, there is also another possibility that we could donate some of this kinds of study at the plants also. That's maybe another thing we could discuss with the utilities. But I hope we can hope to some of this discussion in general when we come to this EPRI user meeting.

MR. BRAARUD: When we run these scenarios, the reason for that we running this within subject design is, one reason is that there are few crews available. So this is the feasible way of doing it.

It has several advantages, but there some also, let's say, issues that we need to consider. And one is the learning effect.

If you run scenario variants, the same crew, they will after some runs, they will be prepared and recognize what is the problem in this scenario. And, of course, this is not a feature we would like to see in the results.

So the scenarios, they are typically what we call counter-balanced so that the different crews, they run the scenarios in different order. And we also make some, let's say, actions or things to hide that they are actually running the same scenario. Like having a small alarm or some small problem early in the scenario that are not important for the rest of the scenario. But just to try to make the crew not recognizing the scenario.

And it's also such that we have balanced the scenario such that they don't run the same main scenario on the same day. Typically they run different variance on different days.

And we try to mix the scenarios so the scenarios have the same, you could say, starting event but have different development. So that we try as much as possible to not have them learn the scenario.

So this is also a methodological choice. If you want to have much data, you risk some learning effect. The alternative is to run much less runs, each crew for example run only one scenario. So this is some choice one have to consider.

And there's also several typically used measures and data collections for the experiments. It's like the reactor operator and turbine operator have a small head mounted camera, the size of a pen, attached to the head to see what information they are looking for in the interfaces, to have a good record of that.

Also all their interactions with the interface are recorded in a log. So you can see when each operator selected a process performance and when they did a action.

There's also some cameras capturing the whole control room. And as I also said, we record all the communication. They have a small microphone attached to each operator. And also all the process parameters or all the important process parameters are logged during the simulation.

And also we have typically a subject matter expert commenting on line when running the scenario. That is very helpful for later analysis to while the scenario is running, actually point to important points in the scenario where we should analyze further. For example, if they did some unexpected action or did not -- or it seems like they did not detect or understand the scenario as we had expected.

CHAIRMAN APOSTOLAKIS: The commentary was done separately, right?

MR. BRAARUD: Yes. This commentary is in a gallery.

CHAIRMAN APOSTOLAKIS: Okay.

MR. BRAARUD: And the crew do not hear these comments. But the commentor hear all the communications of the control room crew.

And also use several questionnaires. For example asking them about the factors that we have manipulated, how did they feel, what kind of time pressure did they feel in the scenarios.

We ask them about the typical performance rating factors. Did they experience any problems with the procedures, any problem with the interface, for example.

We also have some online evaluations.

And also this can differ between different studies, but typically we have the crew to do a debriefing after each run. In this case, it was a debriefing that the crew did themselves, supervisor actually was leading the debriefing.

Then to some results from this experiment. One example, and that will be from the masking research question. The research questions, they are a little bit more specific for each element. I will not use much time on that here. But this is how to see how the complexity of a second or a secondary task effect on the performance of a main task. In this case, it was a relatively simple main task.

(Whereupon, at 5:00 p.m.the meeting proceeded into the evening session.)

E-V-E-N-I-N-G S-E-S-S-S-O-N

5:00 p.m.

MR. BRAARUD:

So the design as described is for variants of a base scenario. It's a main task with additional tasks. So each scenario variant has the same main task, but the variants have different added additional tasks. This is so that all of this scenario variance, they have a leakage from the shutdown cooling system. And this is the main task repeated in all scenarios. This leakage actuate an automatic isolation of the system. And there is two valves that do not close as they should from this automatic system orders. These are two containment valves. And this mean that the leakage is not isolated.

And we have assessed that this main task, we expected to be an easy task for the crew. They have clear indications, they have alarms and temperature in the room where they have the leakage. They have a very clear indication that this automatic isolation have been activated. And they have guidance from procedures. And the action they are to perform when they have decided that this is the case, is a very easy action to perform in the interface.

The additional task is a leakage from the stream pressure relief system. And this is such that in the first variant we used the term "base case" to say that this is a more nominal scenario. There is actually no additional task, there is only the main task.

In the variant number two, scenario number two there is a steam pressure relief valve, a main valve that is faulty open but missing the open indication.

The third variant is a little bit more difficult. There is actually a leakage also through the steam pressure relief system, through the leakage is through one part giving indications in another part. I will actually show a little bit explanation.

The variant number four is the same as number three, with even one more information piece missing.

Just to show one example. This is a process format where they will find that they have two containment valves open. They are in the red circle.

MEMBER BONACA: So this is one of the displays?

MR. BRAARUD: This is one of the displays that the --

MEMBER BONACA: Of course they have no

circle and arrow, but that's -- okay.

MR. BRAARUD: Yes, that's true. Without the circle and without the arrow.

MEMBER BONACA: Okay.

MR. BRAARUD: Yes, that would be too easy for them. That's true.

MEMBER BONACA: But that's from a display?

MR. BRAARUD: Yes, this is from a display. Of course, they have -- in other information, they have the alarms, they have the OE information indicating that they have this isolation activated. But when they have decided, gone through the procedures, that this is the case, they will go to this format and close one of the valves in the red circle. That will actually close the leakage, isolate the leakage for them.

Additional task, this is a breakout from a format from the steam pressure relief system. They have four different, if you can call it, trains or subsystems. And in the red circle there's an indication of, maybe I can just pointer. This is a main relief valve.

CHAIRMAN APOSTOLAKIS: No, you can't do that. Do we have an electronic pointer?

MR. BRAARUD: Maybe I can use the mouse.

CHAIRMAN APOSTOLAKIS: The mouse, yes.

MR. BRAARUD: Oh, yes. Yes. That's perfect. Yes, that's good.

This one is the main valve and this is actually open. This should have had a red indication like this indicating that it's actually open. And they also have indications on temperatures going to the parts.

So this is added in scenario version 2 as an additional task.

In the scenario version number 3 there is actually, if you look into the red circle, this is a more typical example of a mask situation. This is more difficult. The cases that are here through this valve, they have a leakage. This is all the steam pressure relief system. They have a leakage through this valve. But the instrumentation of this plant is so that the steam coming through here will actually activate the indication for this valve. So they have an indication that this main valve is open but it is not, the leakage is through this valve.

They have a temperature indication here indicating that there is something going through this pipe. And they have, let's say, the normal indication that this valve is open.

While the version number 4 is exactly the same, but they even miss this, quite important. They missed this temperature indication for this valve. Okay.

So just to jump directly to some results from this scenario. This is research for the main task. Here we have the four scenario variants, one, two, three, four. This is the time for closing the main task or closing the leakage from the shutdown cooling system in minutes after the leakage was initiated. And we have the seven crews, there are seven staples here, which is the crews named A,B,C,D,E,F,G.

So -- yes. These are actually the performance indication used on the main task, time closing leakage. And this actually mean 20 minutes mean that one crew did not close the main task leakage before we ended the simulation. That was ended 20 minutes after.

Okay. Before we look more at the results, we can also look at the additional task. This is the same type of figure. We have the scenario versions, one, two, three, four. In version number 1 there was no additional task so there is no results. And version number 2 it's the same, it's the minutes taken to close the steam pressure relief leakage. And for scenario 3 and 4, this means that only one crew closed.

MEMBER BONACA: And the same crew closed it?

MR. BRAARUD: And also the same crew closed the leakage.

MEMBER BONACA: Also the crew, they performed extremely well before?

MR. BRAARUD: Yes.

MEMBER BONACA: So there is something special about crew B?

MR. BRAARUD: For this scenario they performed very well.

MEMBER BONACA: Yes.

MR. BRAARUD: That's true.

MEMBER KRESS: What information did they use to decide that the leakage is coming through that -- because it doesn't look to me like they have any. In the fourth scenario.

MR. BRAARUD: Yes. In the version number 3 they have one temperature indication.

MEMBER KRESS: They have temperature there.

MR. BRAARUD: Yes.

MEMBER KRESS: In 4 they had nothing.

MR. BRAARUD: No. In 4 they have actually to -- they would have to infer or try to test where could the leakage be.

MEMBER KRESS: I see.

MR. BRAARUD: But they even had some more information available, but they had to look in the alarm system actually to find some information about this temperature. That was not that easily accessible. But they could have found some more information even.

But putting these two figures together is actually how we looked upon how this different context, which was the additional task, effected their response on the main task.

MEMBER KRESS: How do you quantify that?

MR. BRAARUD: If we want, we can actually quantify this by using some analysis. We call it variants analysis. It actually look upon if there are more variants within the different experimental conditions. But at this stage there are also few crews, so few data that we are not actually looking for quantitative analysis at this point. It's much more define the qualitatively what are the driver of human performance or crew performance.

CHAIRMAN APOSTOLAKIS: Why did crew B perform so well in scenario 4?

MEMBER KRESS: That's an interesting question.

CHAIRMAN APOSTOLAKIS: And everybody else was lost?

MEMBER BONACA: Really, they performed well in all scenarios. In fact, from the slide number 1 --

CHAIRMAN APOSTOLAKIS: Yes.

MEMBER BONACA: Okay. And in the previous scenarios.

CHAIRMAN APOSTOLAKIS: What was B?

MR. BRAARUD: Yes. But this is a very important question. And this is also things we have looked at.

CHAIRMAN APOSTOLAKIS: You have or have not?

MR. BRAARUD: We have. We have. We have looked at.

CHAIRMAN APOSTOLAKIS: So you understand why?

MR. BRAARUD: Yes, we have some -- we called it -- we do some qualitative analysis --

CHAIRMAN APOSTOLAKIS: So you're going to tell us?

MR. BRAARUD: Yes, we'll tell you. We'll tell you.

CHAIRMAN APOSTOLAKIS: Okay.

MEMBER BONACA: So you have to go to cognitive analysis?

MR. BRAARUD: Typically we do an analysis of the communications within the crew and also based on the observations done during the simulations. And also analysis of the --

CHAIRMAN APOSTOLAKIS: Are you spending all the time on just this case. Because I see you have many slides?

MR. BRAARUD: I think we have thought if we present, this is an example, this will illustrate all the methodology. All the scenarios and all the questions are studied by similar method.

CHAIRMAN APOSTOLAKIS: Okay. So we can after you finish, stop there you think? I'm trying to figure out whether we need a break or not. You have ten more slides and then you have time pressure -- oh, no, sorry. The whole thing is this case, right?

MR. BRAARUD: No. That's after about ten slides, we are --

CHAIRMAN APOSTOLAKIS: You are moving to

another --

MR. BRAARUD: Yes, another question.

CHAIRMAN APOSTOLAKIS: But within the same experiments?

MR. BRAARUD: The same experiment.

MR. BYE: But looking at time pressure at the information --

CHAIRMAN APOSTOLAKIS: Well, I don't know. What do you think? Shall we finish this part and then take a short break.

MR. BYE: It depends how long you are going to continue. Because -- it's up to you.

CHAIRMAN APOSTOLAKIS: Any advice? Some we take ten minutes now or continue?

MEMBER BONACA: Let's take ten minutes.

CHAIRMAN APOSTOLAKIS: Let's go on. Okay. Let's go on.

MR. BRAARUD: Okay. Typically when we are comparing the conditions, the scenario variants 1, 2, 3, 4 give us some indication that in scenario variant 3 and 4 there is some longer response times on the main task than on variant 1 and 2.

There was one long response time in scenario variant 1 which was unexpected. And typically what we do, we do what we call a special or a case analysis of those instances where we think there are some important things to look at. And for this first variant, the first crew that I'm pointing to here, it was actually a misunderstanding by the reactor operator in the interface, actually choose the wrong valve first. And after some time he realized that he had not actually closed the leakage. So he closed it. So this was not actually related to if they had an additional task or not. It's an interface issue.

So somehow we say that we can disregard this one.

Some of the other interesting cases is those with long response time. Why do they actually have such long response time, and it could be as a pointer crew B, why do they perform so well and why are they the only crew that solved difficult additional task, scenario variant 3 and 4.

And typically we do a case analysis based on crew communication and make an interpretation, typically a team of several people some with operational experience, some with more human factors psychology experience. And if you look at those crews that have long response times, the reason for the long response time on the easy task is actually that they are occupied with this additional complicated task.

And it's typically that both the supervisor and the reactor operator, they focus on this time. They have problems solving this time. And first, typically the reactor operator try to close this additional leakage, can't do it. The supervisor has to assist the reactor operator. And they actually forget to take the full overview of the plant and the alternative was to actually divide the tasks better within the crew so that one operator work with additional task. And the supervisor, for example, assist in solving the main task, for example.

So case analysis show that the reason for that related to the main task is that they are using undue resources on this problem, the additional problem.

Also if you look at the scenario version number 2 there is some differences in how they solve the additional task. At case analysis we'll give insights to why do they have these differences. And it actually it shows that three of the crews, they make what you can call a correct diagnoses right away. They conclude that the main valve is faulty open and they close it.

While the other crews, they actually make a -- you can say a wrong diagnosis of the situation of this additional task. First, just take one example, that they conclude that the main valve have been opened but it's now actually closed as it is indicated. So they conclude that this is not a problem at this time. But as the scenario run they will have process indications that there is something wrong with the pressure relief system. They have actually effects on the process. For example, the condenser and the suppression pool temperature would be effected by this. But based on this indication from the process, they reevaluate their first interpretation and make the correct diagnosis.

So without going into detail for each crew, this is actually the path done.

So the conclusions from this type of case analysis is that there is actually some variability in how crews, in this case 7 crews, interpret what we would say was somewhat or a little ambiguous process picture. And actually this lead that they make the wrong diagnosis, but all the crews they manage to get the correct diagnosis indicating that they are actually able to recover from a wrong diagnosis as long as they have process indications that point them to that this is not the correct diagnosis of this task.

So this also may be a little bit related to what was discussed previously today. For example, this confusion matrix and the results from that paper pointing to that, let's say, errors of commission which is related to diagnosis. That was the big problem. And, actually, this confirmed this when we talk about quite simple scenarios. I think that could be the case; that this is not a very difficult scenario. They have good indications that they are not on the right diagnosis. And if they get indications to reevaluate the diagnosis, the crew actually performed the correct diagnosis in this scenario.

So this is one type of result from this kind of case analysis.

And also if you look at some example, crew B was mentioned as a very good crew in this scenario. As an example, we have used the scenario variant number 3 where they performed well on the main task and also are the only crew that solved the complicated task. And the case is that it looks like it is team management or delegation of work within the group is one important element. The case is that the supervisor, he notices that the reactor operator is occupied with the steam pressure relief problem, but they have noticed that they have actuated an isolation system for the main task. So the supervisor, he let the reactor operator work with the one task while he himself take an overview of what we call the main task and quite easily close the leakage by closing the valve. And this gives both the reactor operator and the supervisor time to work with the complicated task. So they discussed this task.

And also there is one important instance and the reactor operator, he detects the temperature indication from the pipe where it was actually leaking. Maybe I should just briefly -- the reactor operator he look at the process format for this system and he detect this alarm indication. But he do not actually know the implication of this information. But he communicated to the supervisor that there is an indication or something in with this pipe, which is a very good feature and not all operators in all situations would communicate an information piece that they actually don't have understood fully or know the significance of. So that is what they do in this situation. And based on this information the supervisor actually reasons that this must indicate that there is something going through this pipe. And he make the diagnosis that the leakage could be from this pipe, the correct pipe. And they try to close the valve that would close the pipe. And they actually close this complicated task.

So there is some -- we can call it a characteristic of the crew that they have very efficient team management divided between these two tasks. And they have, let's say we can call it very open communication or it is allowed to communicate the piece of information that the reactor operator actually is not sure about the meaning of, but he reported to the crew.

So this is also more insights.

So my conclusions from this masking scenario is that for the version number 2, which was not a very difficult additional task, four crews actually made what we can call an initial wrong or incomplete diagnosis of the additional task. But this had no adverse effect on the main task, actually. They were able to solve both the main task and the additional task at reasonable times.

But when having more difficult additional tasks, this made the crew using resources on this complicated task not solve that task. And that actually resulted in reverse response for several other crews of this main task. That can effect all the context on the same quite simple main task.

MEMBER BONACA: The question I have is the masking process resulted in, for example, temperature variations in the display that they had. But what kind of symptom did they have in control functions? I mean, did the masking also effect the transient of the main event that they were simulating?

MR. BRAARUD: No.

MEMBER BONACA: It didn't.

MR. BRAARUD: Actually, the additional task leakage did not effect the leakage from the shutdown cooling system.

MEMBER BONACA: Okay.

MR. BRAARUD: You could say they were independent. After they had this leakage manifested, they were independent of each other.

MEMBER BONACA: So how did they know that they had a masking event?

MR. BRAARUD: Actually, they did not know that they had the masking event as such. They actually experienced that this was a difficult task for them to solve.

MEMBER BONACA: Okay. So they were looking at the displays but they really did not know that there was a leakage there and there was no way that they could understand it from -- it would be only from the temperature variation. I'm trying to understand how they would look for it.

MR. BRAARUD: Yes. Okay. Yes. They have indications that there is a leakage in the steam relief system.

MEMBER BONACA: Okay.

MR. BRAARUD: So that will be manifest in some of the main process parameters.

MEMBER BONACA: Okay.

MR. BRAARUD: So they know they have a leakage there, but they are not able to find the cause.

MEMBER BONACA: Okay.

MR. BRAARUD: Yes.

MEMBER BONACA: So they really had clues from the process parameters --

MR. BRAARUD: Yes.

MEMBER BONACA: -- and -- okay.

MR. BRAARUD: So based on these results, the case analysis -- because general conclusions where I set up in the report, actually describing how the context effected the main task. So it summarizes some of the things that I said here. That the secondary task has the potential to effect the performance of an easy main task. If there is some indication or resulting process deviations that are indicating that they have a secondary task and not only the main task. And they have to judge or they have to prioritize to work with the additional task if this is going to effect the main task.

CHAIRMAN APOSTOLAKIS: Did anyone of the models we heard about today refer to multiple tasks or do they all focus on one task?

DR. ELAWAR: They do refer -- the original -- the original is used much higher from the third table if you have a high workload or, so to speak, more than one task going on and the stress factor as well goes up. If you have a second event within an event, it will go to a different third table and it may lead to a higher stress factor.

CHAIRMAN APOSTOLAKIS: But this is not what these guys are talking about. They are talking about misdiagnosis, different --

DR. GERTMAN:: Excuse me. Dave Gertman, just for the record.

In SPAR-H what we do is we'd increase a PSF for complexity and probably stress, we take a look. That's how we manifest the introduction of a second task as to complexity and part of that diagram on that particular PSF.

CHAIRMAN APOSTOLAKIS: Well, but this is really a very interesting result. And I guess part of the qualitative analysis or insights that these gentlemen are talking about is exactly that; to figure out how does my model handle this, right? Now, if you handle it, you handle it. I mean, I'm not saying that you're not.

I think in an earlier slide that the tasks -- yes. Slide 45, the previous one.

MR. BRAARUD: The previous one.

CHAIRMAN APOSTOLAKIS: The key word here is "easy main task." So --

MR. BRAARUD: Maybe it could have been that it has the potential to effect even an easy main task.

CHAIRMAN APOSTOLAKIS: Yes. The models we're talking about here will start with the easy main task, assign a probability and then they will go to the secondary task and assign an conditional probability. That's not what this says. This says that the performance even in the first task, which was declared easy, is effected by this second task.

MR. BRAARUD: Sure.

CHAIRMAN APOSTOLAKIS: And I think it's a subtlety that perhaps we should confront.

MR. JULIUS: In the EPRI approach it would come through in two cases. With the cause-based decision tree there's a specific failure mode for failure of attention and it's driven by the low and the high workload and the complexity. So you would see even for the first task if there's a high workload, that the probability would be effected. And the cognitive response would be the impact on the response time. You'd see that with the additional complexity in the masking that the response times would be longer.

CHAIRMAN APOSTOLAKIS: Okay.

MR. BRAARUD: Okay. So there are also some properties of this secondary task, maybe I don't have to repeat them, but --

CHAIRMAN APOSTOLAKIS: No.

MR. BRAARUD: Yes. So also what you see is actually that there is an interplay between the, you can say, the process driven context and the preparedness of the crew. So typically if there are weaknesses in how the crew work, for example resource allocation, this complicating scenario driven by the process will become manifest as a problem if this two features or maybe you can call them PSFs or factors are brought together.

So it also point to those crews that have very efficient resource allocation and efficient supervisor managing the team. They are more able to handle these kind of scenarios.

CHAIRMAN APOSTOLAKIS: Well, there is a risk here of getting lost in the details, though.

MR. BRAARUD: Sure.

CHAIRMAN APOSTOLAKIS: Because, you know, you're running these experiments, you have all this information, you know you reach a nice conclusion. Now when you start getting into resources and this and that, remember that in the PRA the numbers are really low. I mean, they're covering a broad range of impacts. So the interest is -- I'm not saying don't do this. But what I'm saying is the interest really from the PRA perspective or the HRA perspective is have we captured the essence of this, not whether are undue resources weren't here or there. Because I'm sure in every scenario you will have a lot of observations that probably are grouped in a PRA. I mean, we're not doing such a detailed analysis that allows us to account for every single thing. But, again, I'm not saying don't do it because these are important things.

Okay. Let's go on, unless there are questions.

Guys, we have to make a decision here. There are 20 more slides. Either we take a break or we ask these gentlemen to jump into conclusions. What do you prefer? Mario? I think we should go over all the slides.

MEMBER BONACA: Yes, I think so, too.

CHAIRMAN APOSTOLAKIS: Well, let's stop for a while.

MEMBER BONACA: If one needs a break, then they can get up.

CHAIRMAN APOSTOLAKIS: Well, they can't do that. The reporter can't do that. So let's take ten minutes. It's still early. Okay.

(Whereupon, at 5:31 p.m. a recess until 5:45 p.m.)

MR. BRAARUD: Okay. Should I start again?

Shall I try to make it a little bit quicker for the more examples so we just get a feel of--

CHAIRMAN APOSTOLAKIS: Yes.

MR. BRAARUD: I'm not going to show a little bit about another part of the experiment that focused on two other dimensions, the time pressure and the information load part. And in this case there is also a main scenario which was actually an incomplete scram scenario where they have to -- in the end they have to start the boron system.

MEMBER BONACA: Yes, pretty slick.

MR. BRAARUD: Yes. So there is control routes that are stuck and also some scram valves that are going to scram valves that are going to shot the rods into core that do not open. Okay. So this is the main task. The most important task is to start the boron system. There is also some other additional task, but that's an important one.

And there are some additional tasks set that was expected to create more time pressure for the crew. And there is, in this case also, a main steam pressure leak system valve that is open. And there is also the initiating event to this scenario was that they have problems with the feedwater and they have a feedwater isolation.

MEMBER BONACA: Wouldn't these be in masking effects, too? I mean, they intended time pressure, but they're similar to the masking scenarios you had before, are they?

MR. BRAARUD: They are. But the indication in this case on the main steam pressure relief valve are normal. So there are no -- planned problems, additional problems with this task.

So they have this steam pressure relief valve open. They have also some auxiliary feedwater trains that are not working as they should and they need to work also with these trains. And there is some tasks that we expected to create more information load, there is some decreasing level in the feedwater tank. They have some alarms on the intermediate cooling system. They have some vibrations on one on the recirc reactor, recirculation pumps.

It's the same in this case, actually, we have a base case which is the main task only. Scenario variant number 2 we added the task expected to create time pressure. Number 3 we added the task we expected to create information load. The fourth variant we added all the additional tasks, both those to create time pressure and information load.

So the fourth variant should be seen as the most complicated context for the main task.

A table showing some of the main results. You have the scenario variant in the rows. Okay. It's only internal number. This was scenario number 4. But you have the .1, 2, 3 4 here indicating the variants. And you have the crews. And this is the response time in minutes when they started the boron system after the incomplete scram.

MEMBER BONACA: The crews are the same that you had before. So crew B was the one that was very successful before?

MR. BRAARUD: Yes, these are the same crews. B is the same crew as before.

Only in this case the performance indications used was actually how much the crews deviated from the mean. So in this case also we can see that version 2 there is one crew with a long response time. Number 4 there is three instances with longer than one standard deviation from the mean. And there is also some indication it was estimated, actually, based on the task the crew needed to do and the procedures that the nominal time to perform, start the boron system, was 12 minutes. That is in expert judgment material. It's not from any technical specification or anything.

But I guess a training instructor would expect to do that in five minutes.

So this also indicate crews with slightly more time, lower response time than the --

CHAIRMAN APOSTOLAKIS: So it's five minutes for all scenarios, the expected time?

MR. BRAARUD: Yes. Yes. Yes.

CHAIRMAN APOSTOLAKIS: Even though you added things?

MR. BRAARUD: Yes. It is expected that these additional tasks, they are actually quite quick to solve if they handle the tasks correctly. Of course, there is a very minor difference in the nominal time, you can say. But we expected it to be the same.

MEMBER BONACA: The main task?

MR. BRAARUD: The main task.

MEMBER BONACA: The main task, did they accomplish all, I mean within the five minutes?

MR. BRAARUD: No.

MEMBER BONACA: No, no, the main task? Oh, the main task.

CHAIRMAN APOSTOLAKIS: Scenario 1 is the main task, isn't it? 4.1 is the main task?

MR. BRAARUD: Yes. Actually two crews used also longer time than the nominal time, which was also a little bit unexpected.

MEMBER BONACA: Yes.

MR. BRAARUD: But as you can see, those with the longest times, they are in the variants with either the time pressure only, one crew took a long time, or in the version with both time pressure and the information --

CHAIRMAN APOSTOLAKIS: But it's interesting, though, that crew G --

MR. BRAARUD: Yes.

CHAIRMAN APOSTOLAKIS: -- performed better when you had both time pressure and something, information load was the other one?

MR. BRAARUD: Yes.

CHAIRMAN APOSTOLAKIS: In 4.2 they didn't do so well. Presumably, 4.2 is simpler than 4.4?

MR. BRAARUD: Yes. Yes.

CHAIRMAN APOSTOLAKIS: So what was going on there?

MR. BRAARUD: Here we also have some instances are learning effects through the scenarios. So by running one crew through several scenario variants, there will be some learning effects. These learning effects we try to spread out in the data set by having all crews running different orders. So it is important to look at the pattern of all the runs.

And also there are many other factors or many things that can effect the performance of the crew. It is actually not the case that a crew that runs several similar scenarios, they do not actually perform the scenario the same way the subsequent runs. There are minor variations that will create -- that is actually the way human operators are. They are not that consistent that we --

CHAIRMAN APOSTOLAKIS: So there's aleatory effects here?

MR. BRAARUD: Yes. So there is some minor effects so that they actually choose to work a little bit different. They use a little bit more time on the procedure. They were actually looking at some other process format than the previous run when the event came up. So, some minor variations will always be in the data. So we look for that, the pattern.

CHAIRMAN APOSTOLAKIS: Well, 11 minutes is not minor.

MR. BRAARUD: No, that's a long -- and also here can also see that in this case crew G has two other long response times, and also the longest ones.

MEMBER BONACA: Of course they need to perform the main task --

CHAIRMAN APOSTOLAKIS: Three cases they are the longest.

MR. BRAARUD: Yes. So this actually indicate that this crew G represent some, let's say, characteristics or potential that this scenario variants actuated and maybe they have not --

CHAIRMAN APOSTOLAKIS: Maybe they were not experience, is that possible?

MEMBER BONACA: Homer was a member --

CHAIRMAN APOSTOLAKIS: What?

MEMBER BONACA: Homes was a member of this crew.

CHAIRMAN APOSTOLAKIS: Maybe they were not as experienced as the other crews?

MR. BRAARUD: They were experienced.

CHAIRMAN APOSTOLAKIS: They were experienced?

MR. BRAARUD: There was not different from the mean, actually.

CHAIRMAN APOSTOLAKIS: Interesting.

MR. BRAARUD: Also some of the insights we can have from this run is that not always only experience that is important for their performance. That can be for some scenarios important, but not for all. Because many of these crews, they have passed a -- because they are -- they are very good trained, generally. So even you have three years experience as a supervisor, you can actually perform in many instances as well as one with ten years.

MEMBER BONACA: The interesting thing, in crew D, crew D actually did much better when there were additional time loads and things of that kind.

MR. BRAARUD: Yes. Yes.

CHAIRMAN APOSTOLAKIS: Well, look at crew A, they did their best in the most complex scenario.

MR. BRAARUD: Yes.

MEMBER BONACA: Again, in that order, and that's what I'm looking at.

MR. BRAARUD: This likely has to do with the order effects, some learning effects and you can say some random variants. But this also indicate that it is not very strong effects of this time pressure, but there is some effect that we can see when we look at the whole data set.

So in this case also we can do similar types of analysis that we did for the previous scenario going into detail why did some crew perform good, why did some crew have problems. And for the performance there also, the scenario this time is quite similar. There is an additional problem in the scenario. And those crews, crew D that performed not that good, it's related to the same phenomena, that they actually don't manage the resources as well as the teams that perform well in all conditions.

So this seems to be for this kind of scenarios with several task, this teamwork management or this type of crew characteristic, it's important.

I don't know if PRA actually take into account what kind of training do the population, let's say the sector operators have for the plant. Do they have, for example, specific training at handling multiple tasks, for example. Would that mean that they would perform better than a plant that don't have this kind of training.

CHAIRMAN APOSTOLAKIS: In principle it should be taken into account. I don't know whether in practice we actually do that. I mean, to declare a crew as novices is not something that's easily done. Because it's done on the average. When you do a PRA, you don't have a particular crew in mind.

DR. ELAWAR: Correct.

MR. BRAARUD: Yes.

CHAIRMAN APOSTOLAKIS: I mean in books you see things that say, you know, adjust it if it's novices and so on. But in practice, I'm not sure how much --

DR. ELAWAR: In practice we are still trained operators.

CHAIRMAN APOSTOLAKIS: Trained operators.

MR. BRAARUD: Okay. So this is some more detailed results actually showing how they performed on the additional task and that this relate to how they performed on the main task for some of the runs. But actually not for all of them. So there is some explanations of why it was related to some of the crews and why not.

CHAIRMAN APOSTOLAKIS: So what does that no mean?

MR. BRAARUD: No mean that they did not actually close the steam pressure relief leakage. In this scenario it is quite complicated logic in the system. They have some, what you call it, interlocks or preconditions that they can only have one valve open in a given train. And they have to close one valve that is already open to be allowed to close to another isolation valve that actually close the leakage. And this is something that some of the crews had problems with in the scenario.

And there are some case analysis. For example, explaining why one crew have a very long response time. And they're just taking from some of the transcripts of the scenario.

So actually there is instances that they have a problem with this additional task. And also some instances that they don't communicate quite good. There are also some nearly disagreement between the reactor operator and the supervisor what is actually the best approach. And they are both experienced people. So there are some issues explaining the long response time.

CHAIRMAN APOSTOLAKIS: What was your role in this? You were just observing?

MR. BRAARUD: My role in this experiment is typically we conduct the family experiment, specify what items should be researched, making the research plan. And also we have participating in collecting the data. There is quite a big, call it organizational work to run all these crews through all the scenarios, collecting all the data. And we also do the analysis, there are several people involved who perform the case analysis and the conclusions.

MR. BYE: Maybe we should mention that there are also experts joining to decide the scenario. And has worked between 10 and 20 years as supervisors and operators in Sweden actually.

MR. BRAARUD: Yes, so it's a team with several competencies.

Yes, so this team management division work turned as important. There are some more analysis.

In this case for this scenario, those additional tasks that we had defined as information load, they were actually not a problem for the crew in this case. And this may have to do with the characteristics of these additional tasks. They were say, correctly, considered as not important and not prioritized to work with.

So tasked with these characteristics, created no problem. Was no problematic additional context for the crew.

We, again, take one more scenario briefly where we also studied time pressure and the information load factors. And the event in this case was a loss of the main grid, external grid, which for this plant resulting that they produce power for their own use. They call it the house turbine operation. And they have a backup grid available. And the procedures say that they should transfer their -- or get the supply from the backup grid. And this has to be done manually.

CHAIRMAN APOSTOLAKIS: Is that automatic in American plants?

PARTICIPANT: (Off microphone).

MR. BRAARUD: The case is that the transfer itself for this plant is automatic. But it has to be started manually. So there is an -- I don't know how this works. But there is automatic sequence that will transfer, synchronize and transfer this. But the operators have to decide that they will do this and manually start it.

CHAIRMAN APOSTOLAKIS: I wonder why it's not automatic?

MR. BRAARUD: There may be reasons. I cannot tell you that.

But the case is that they have a air leakage also in the turbine condenser that will give them a trip of the turbine. And this will actually if they don't have transferred to the backup grid before this trip, this will actually give them a scram and they will automatically start the emergency power supply.

CHAIRMAN APOSTOLAKIS: Does the reactor scram when they lose outside power?

MR. BRAARUD: No, they have no reactor scram. They have a reduction in power. It's regulated down to 50 percent, I guess. But they run the plant to produce enough power to support -- to supply the plant. So that's why this plant is designed that way. So they do have a reactor scram when they are in this situation.

Okay. So they have some advantages of transfer to the backup grid. They will not have the emergency backup, the emergency power starting up. There are some sequences that will actually stop several important components in the restarting to not overload the power supply.

Also they have four trains that should be manually transferred and there can be different arguments to transfer the different trains. I don't know we have a slide. But the time pressure in this case is also that they have a leakage from the steam pressure relief system, but the time pressure is so that they will have a reactor scram earlier in the scenario.

In the base case they will have, let's say, 25 minutes when they have this leakage. The time pressure case, they will have shorter time. Maybe 15 minutes. I don't remember exactly. It's in the report, but around there.

And there's some also some information load tasks, which was we expected them to use some time on this task, but diagnose or prioritize so that they don't need to take this task into consideration.

Okay. Jumping directly to the results.

I didn't say that much though that they

had four trains and there were different arguments for which trains they should transfer. This is related to which components is supplied by the different trains.

CHAIRMAN APOSTOLAKIS: I don't understand that. What do you mean arguments? It's not part of the procedures?

MR. BRAARUD: No.

CHAIRMAN APOSTOLAKIS: They have to decide?

MR. BRAARUD: They have to decide the order.

CHAIRMAN APOSTOLAKIS: Why? Shouldn't it be in the procedures?

MR. BRAARUD: I think that -- I'm guessing a little bit, but I think that the procedure is written for a situation where they don't have any problems or reason to prioritize. Maybe they may have some reasons that, let's say one of the trains supply important components like feedwater, for example, maybe. But as I have heard there is no priority given in the procedure. It's actually stated they should transfer these four to the backup.

So there is also some issues of why did they prioritize the different trains, what kind of reasoning did they actually use; that's one part of it. But also it is interesting to see that in the time pressure scenario, I will not explain this table it's a little bit detailed, but these shaded areas mean that they performed the scram without any transferring of those parts, meaning that they will then rely on the emergency power supply. They can later commit it to the backup grid, but there will be, you can say -- yes, they will actually have some components without power for some period.

So a training instructor would say this is not idle or not tested even though the expected solution. Most crews do not do it.

So in this case the context, what we thought to be a time pressure task, seems to be the course for two crews actually feeling that they needed to scram the reactor. They didn't have enough time to perform the transfer or they actually considered it and more important to scram the reactor due to this steam pressure leakage than to perform the -- no. One crew actually deliberately discussed if they should do it or not. Three other crews,they actually more, I will say, forgot the transfer problem and decided that the most important thing is to scram the reactor in this situation.

And for similar reason in variant 4 also one crew started to actually transfer one without succeeding and scram the reactor.

So the case that one crew within the information load scenario so performed, scram the reactor, that was actually based on they both acted -- I will say, a fault with the simulation actually that they had some oscillating steam valves, that they considered to indicate some oscillations in the core and they decide to scram the reactor. But this was, I think, the most interesting result from this scenario. I'll just jump to it. It's actually that the added task that we thought should be an information load task was actually integrated by several of the crews as time pressure. They used actually the same amount of time before transferring to the backup power as in that scenario with time pressure. So when we were analyzing these scenarios we were thinking that this temperature was actually just passing a level and they should actually not consider as this an important task that should actually make them feel that they had in this case tripped the reactor.

The same with a vibration alarm on the turbine bearing. It also fluctuated around the level and they thought they should not consider this an important task.

So it shows that the crews, they perceived this situation differently than the analysts doing analysis without running or without experience of running the scenario. So it pointed out there are other things, for example time available that we can calculate more objectively. It is a very good example that if the crew feels that they have time pressure to do an action or they can be that they feel it's important for safety reasons or for equipment, prevailing equipment, they perform this action.

I guess this actually sum up some of the most important research from the experiments, some of them. It give a good indication of the method used on the question studied and how similar experiments could be performed.

MR. BYE: Maybe one thing to this crew, which crew was a good one, we should say that this A,B,C,D,E,F,G numbering is not sequence. This is randomized.

CHAIRMAN APOSTOLAKIS: Okay.

Should I take a little summing up, I think, this HRA and PRA implications. We just have summaries of this. I don't think we have gone through this before, so I don't know if you want to -- yes. Summing a little bit on the implications of how these results can be used in the methods, but I think also this should be left to maybe the people reading this.

So to sum up, we think that this can be used both to inform HRA practices on method development and also giving input to other repositories and so on. So the method is to have controlled study, use scenario variants, look at the external things, but also driving to detail measures and characteristics.

So the next steps. We have been asked to document this methodology to maybe make --

CHAIRMAN APOSTOLAKIS: Then you better do it.

MR. BYE: Of course, this experiment is documented here. We also want to document this also related to the HRA methods and so on, but also to peer review that and to get some feedback on that.

We are going to run more studies in 2006. And we have started on one study, and that is to run one study on our PWRs going further in masking and other PSFs. We have had one crew from the Swedish -- this is a Westinghouse, this is a 900 megawatt Westinghouse two loop plant. One crew so far. But they are doing upgrades and have problems with supporting us with crews. So we would very much like, both from them or from the U.S. crews, to join us in that.

CHAIRMAN APOSTOLAKIS: You heard earlier that there was a disagreement of sorts between Dr. Cooper and Dr. Gertman. Dr. Cooper felt that going with a context was a very important approach, whereas Dr. Gertman said my PSFs cover maybe 80 percent or more of the context. I don't need to go to such a detailed evaluation. It would be very interesting if you could devise experiments that would shed some light on this difference. I mean, I appreciate that you're now looking at individual factors and trying to understand what's happening, but maybe down the line you can figure out something and say this -- I don't know how you do that, of course. You have to plan it. But, you know, in this case it was really context. I don't know how you would do this. But that's why we're running experiments. And that the PSFs in a similar situation appear to capture the whole issue. That would be extremely valuable.

And you are focusing now on time and information load, of course there are other PSFs as well, as you know, in one table. SPAR-H lists eight of them. It would be nice -- and the other thing that is a little bit up in the air is this also the duality of the PSFs. I mean, are they really independent?

It seems to me if you look at the existing PRAs, the various models, and look at the criticism and so on, maybe the document the NRC is preparing comparing these models to the Best Practices document, there's a lot of useful comment area there. And maybe you can look at it and try to see whether one could device experiments that would, again, shed light on these controversies. That would be very useful.

MR. BYE: We have been discussing this, or not benchmarking maybe, but some kind of comparison or looking into methods by maybe running sequences, classifying them and then running them in a lab.

We discussed this in an HRA workshop in Holland one month ago with several people from other actuary method developers also in Europe. There is some mixed motivation for doing that. And I think we need really to go into a cooperative effort with very many method developers to do that in a way that really can be accepted by --

CHAIRMAN APOSTOLAKIS: Well, I'm not saying straightforward. But I mean these are the issues that seem to be sort of unresolved regarding the models. Also, if you can shed some light on the various adjustment factors that if time pressure is high, a factor of 5 is reasonable, or a factor of 2 is reasonable. You know, that kind of stuff. Because we're going to have to live with those for a while.

Any questions, other questions or comments? Members? NRC staff? Members of the public?

We appreciate very much your coming here all the way from Norway. It was a very, very interesting presentation. In fact, I was thinking while you were talking how we can have a presentation to the full Committee on this. Don't you think that would be useful with some informational meeting? We had one from Bruce Holbrook some time ago on similar things. It's quite a while. But maybe a presentation along these lines.

Yes.

DR. LOIS: Halden is going to be here for the -- and we can hold them here for about a month so they can --

CHAIRMAN APOSTOLAKIS: I think the timing is not very good, but if they're willing to stay for three weeks in the United States. Is the NRC paying for all of this? Then take your wives and some vacation.

PARTICIPANT: It comes out of their

general funds.

CHAIRMAN APOSTOLAKIS: Where did you come in here?

PARTICIPANT: I've been here all day, George. Didn't you notice me?

CHAIRMAN APOSTOLAKIS: No.

PARTICIPANT: See, I'm so quiet.

CHAIRMAN APOSTOLAKIS: So thank you very much, gentlemen. This was very, very good. We appreciate it.

And on that happy note, we will recess for the day and tomorrow at 8:30 we'll hear how this stuff issued in Bayesian updates.

(Whereupon, at 6:20 p.m. the meeting was adjourned, to reconvene tomorrow morning at 8:30 a.m.)