Official Transcript of Proceedings

NUCLEAR REGULATORY COMMISSION

Title: Advisory Committee on Reactor Safeguards
Reliability and Probability Risk Assessment
Subcomittee Meeting

Docket Number: (n/a)

Location: Rockville, Maryland

Date: Friday, February 22, 2008

Work Order No.: NRC-2040 Pages 1-305

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)

SUB-COMMITTEE ON RELIABILITY AND PROBABILISTIC RISK ASSESSMENT

+ + + + +

FRIDAY,

FEBRUARY 22, 2008

+ + + + +

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

MEMBERS PRESENT:

GEORGE E. APOSTOLAKIS Chairman

JOHN W. STETKAR ACRS Member

SAID ABDEL-KHALIK ACRS Member

NRC STAFF PRESENT:

ERASMIA LOIS

GARETH PARRY

JOHN MONNINGER

NATHAN SIU

HOSSEIN NOURBAKHSH

ALSO PRESENT:

JOHN FORESTER

VINH H. DANG

SALVATORE MASSAIU

DENNIS BLEY

PER OLVIND BRAARUD

JEFF JULIUS (via telephone)

SUSAN COOPER (via telephone)

T-A-B-L-E O-F C-O-N-T-E-N-T-S

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

8:31 a.m.

OPENING REMARKS AND OBJECTIVES

CHAIR APOSTOLAKIS: The meeting will now come to order. This is a meeting of the Advisory Committee on Reactor Safeguards, Subcommittee on Reliability and Risk Assessment. I am George Apostolakis, Chairman of the Subcommittee. The Subcommittee members in attendance are Said Abdel-Khalik and John Stetkar.

The purpose of this meeting is to discuss the draft report, International HRA Empirical Study, Description of Overall Approach and First Pilot Results From Comparing HRA Methods To Simulator Data. The Subcommittee will hear presentations by and hold discussions with representatives of the NRC staff, Sandia National Laboratories, the Paul Scherrer Institute and the Electric Power Research Institute. The Subcommittee will gather information, analyze relevant issues and facts and formulate proposed positions and actions as appropriate for deliberation by the full committee.

Hossein Nourbakhsh 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* in February 2008. 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 not received any requests from members of the public to make oral statements or written comments.

This Subcommittee and, of course, ACRS itself has been very much interested in human reliability models. We wrote a letter, I believe, last year where we recommended that the staff start working with various stakeholders to develop one method or a suite of methods or models appropriate for various problems and the Commission issued an SRM asking the staff to, in fact, start working on this and to also work with and the staff proposed this empirical study would be the first major step towards this ultimate goal of creating a suite of models for use.

So this is -- We have heard about the planning of this exercise of this study and today we will be presented with some results. Things started happening and we're very pleased also to have visitors from other institutions from Europe. So without further ado, we will turn it over to the staff.

MR. MONNINGER: Good morning, Professor Apostolakis, fellow ACRS members. My name is John Monninger. I'm the Deputy Director for the Division of Risk Analysis in NRC's Office of Nuclear Regulatory Research.

First off, I would like to thank you very much for the opportunity to brief you today and to solicit comments from the ACRS on this ongoing project. As you mentioned, this is a very important project for the staff. There has been considerable interest from the ACRS and considerable interest from the Commission.

One thing I would like to highlight would be as you did mention the notion of involvement with stakeholders. In that regard, this project goes a long way to doing that. Though the NRC is intimately involved in this project, its success is dependent upon multiple international parties including Halden, the Paul Scherrer Institute and various other countries. In addition to that, EPRI is involved in this project and several U.S. utilities.

In terms of the NRC's involvement, we are directly supporting staff, our staff, and Sandia Labs and there is also some efforts at Idaho National Lab. But in terms of leveraging resources I think this project is an excellent example. I mean there are individuals across the world that are being supported and funded by their own organizations to contribute to this project and we are very much appreciative of the opportunity to participate with them. We do not believe that we would be able undertake a project like this without their support and very active involvement.

So with that said, I would turn it back over to Erasmia and John and Gareth and we do really look forward to your comments and insights.

In the back of my mind, as you're aware, every year the NRC Office of Research has to, Erasmia doesn't know this yet, but the NRC Office of Research has to propose topics to the ACRS for a quality review. You know, ultimately I think this project has the potential to make that list. We haven't made any determination yet as to what it will be in a couple of years. It could be potentially two years down the road or so when it wraps up, but I think this may be on our list and so therefore it is very important for us to inform you and to keep you on board with our approach and for us to address your comments.

MS. LOIS: I just want to know that, Jeff Julius from Scientech, representing EPRI may be on the phone and he is going to be the next speaker after me remotely. Is he? Jeff?

CHAIR APOSTOLAKIS: May be or he is?

MS. LOIS: He should be. He should be connected.

CHAIR APOSTOLAKIS: Where is the phone?

MS. LOIS: Have you -- Is the bridge here?

CHAIR APOSTOLAKIS: So it's 5:30 a.m. for the poor guy?

MS. LOIS: Yes.

(Laughter.)

(Off the record comments.)

CHAIR APOSTOLAKIS: I think we can start the presentation.

MS. LOIS: Okay.

CHAIR APOSTOLAKIS: He knows what you're talking about, doesn't he?

MS. LOIS: Jeff does.

CHAIR APOSTOLAKIS: Okay.

MS. LOIS: But he is coming after me and the purpose of my --

CHAIR APOSTOLAKIS: Excuse me. Yes. I'd like to note that ACRS member Dennis Bley is --

(Telephone conference connection.)

MS. LOIS: Hi Susan.

MS. COOPER: Hi Erasmia.

MS. LOIS: We are here. So Susan Cooper is one person who is connected.

CHAIR APOSTOLAKIS: Are we on the record now? We are.

MR. JULIUS: Good morning.

MS. LOIS: Good morning.

CHAIR APOSTOLAKIS: Anybody else?

MR. JULIUS: Jeff Julius.

MS. LOIS: Okay Jeff.

CHAIR APOSTOLAKIS: Okay. Before we start, I would like to note that Dr. Dennis Bley, an ACRS member is present, but today he will not act as a member of the Committee. He has worked on parts of this project. So he's here actually working with the project, not as a member of the Committee. So Dr. Lois.

I. OVERVIEW/MOTIVATION OF STUDY

MS. LOIS: Okay. Well, thank you very much. I just want to note that this study is just one aspect of the Human Reliability Research Program and I recognize new Committee members here and probably on a different briefing we could brief the Committee on all our various aspects of the HRA program.

This is just one program that is research program that as noted before and I also would like to thank our colleagues from OECD Halden and Paul Scherrer Institute that are here today to brief the ACRS on our behalf on what we're doing.

What we're going to do today throughout the day is we're going to present the work that has been performed so far and we'll accomplish that by providing a brief overview of the study, why we do it and what we expect to get out of the study and then we'll get into the methodology, the results of the pilot phase and then what we learned. We'll discuss what we learned from that and how actually we can take advantage even of these pilot results to start applying in improving human reliability. And the main purpose of this briefing is to obtain feedback from the Committee for the continuation of the study.

CHAIR APOSTOLAKIS: I think it would be nice to hear at some point how this particular effort fits into the bigger question that the Commission has raised and the Committee. I mean, are we going to have a model or a suite of models at some point so everybody knows what they do and they agree. So we can close the issue of human reliability and, if so, when?

MS. LOIS: If you recall about a year ago when I --

CHAIR APOSTOLAKIS: You don't have to answer it now. All I'm saying is at some appropriate point I would like that to be addressed.

MS. LOIS: And what I'm trying to say here is that this is just one possible study. Also we have activities, collaborative activities, with EPRI to address the bigger question of the importance of HRA on decision making and actually we don't have results yet of that work.

So there are many kind of efforts that we're trying to address the question. This is just -- I think at the end of today's presentation we'll have some important initial results but not the answer. What we tried --

CHAIR APOSTOLAKIS: Would you -- I mean, do you think it would be appropriate at some point for the subcommittee to be briefed on all these activities? You are doing what with EPRI and others? Would it be the next two or three months appropriate for you to come back here and tell us because ultimately we have to work towards the request or the direction of the Commission in the SRM. So we have to be able to say they are doing this piece and that piece and this other piece and they are all coming together in March 2010 and we're going to solve the issue. So that high level thinking I think is important for us not to forget. Okay.

MS. LOIS: I think it is and we should do that at a different time.

CHAIR APOSTOLAKIS: Okay.

MS. LOIS: Because right now we're really pressed to present with what we've done.

So what we tried to do through this specific study to examine the capability of the methods, to predict crew performance in simulators and through that, identify to examine whether or not the methods are identifying drivers of successes or failures and actually how close they come to estimate human error probabilities.

We believe that the outcomes of these studies will help us to characterize the strengths and weaknesses of the methods, provided technical basis for improving the methods and that it could be that we abandon some of the existing methods and we converge in fewer ones. But this study and some other activities we have in the HRA program will provide the technical basis to make the decision as to which methods are best or how we can create a method or improve that will help us address regulatory applications without having a whole host of different methods, if necessary.

CHAIR APOSTOLAKIS: The identifying weaknesses of methods, it seems to me you're going to have a big problem there. Who is going to decide that? The group? All of you? But you have conflicts of interest. Some of these methods are yours. Are you going to write down ATHEANA has the following weaknesses or is EPRI going to say, "Gee, the calculator is really not a good thing to do."

I think there is a real problem there. When it comes to weaknesses, I don't know how you're going to handle that. I just don't see how a group that has a conflict.

MR. FORESTER: We are certainly working together as a team. But on the other hand, for example, Jeff Julius reviewed the ATHEANA methodology, the results of that analysis.

CHAIR APOSTOLAKIS: Yes, but you will have to agree if he comes up with any negative comments, I mean, before anything shows up in the report. In other words, what I'm saying is that you should be very sensitive to this fact because you may end up with a report like so many we have seen that says, "And this method does this and this other method does that. Thank you very much. Let's love each other."

Well, you know this is too big for that. We're spending a lot of money here. So I don't know who is going to be that ultimate judge.

MR. PARRY: I think there's maybe a different way of looking at it though and that is that to not necessarily say this method is right, this method is wrong, but that is this method good enough to support certain applications and that's certainly the perspective that I think that NRR would have on this.

CHAIR APOSTOLAKIS: But I have seen already that in the report that you already have here. You're concluding with some very nice words that everyone is considering the performance-shaping factors and there may be some differences and then you come out and say what's important and this and this and that. But you are very careful to say that everybody is doing a fine job.

MS. LOIS: You're thinking about the existing --

CHAIR APOSTOLAKIS: Yes. So before this gets out of hand, I think you really have to think about it. I think it's a problem. I'm not saying that you are doing anything wrong. But it's a problem if you have three or four groups each representing their own method. Right? And then you have to decide that method B has a problem. How on earth are you going to be allowed to say that? I don't know how you're going to do that.

Now you may have all kinds of noble words today. But in my experience when it comes down to actually writing it down, that's why this conflict of interest idea was developed. So just think about it.

I'm not saying you should have the answer today and don't tell me that we are all objective scientists. I mean, yes, we are. It is a problem. To me it's an inherent problem of the effort. It is extremely important you guys to work together, but then I don't know how you're going to write your final report.

MS. LOIS: For me, being on the optimistic side all the time, I see that the various stakeholders in the study to be really willing to recognize the potential weaknesses of the method and also with respect to the NRC and industry stakeholders I believe that there is a real interest to move forward recognizing how we can improve the whole host of tools and I wouldn't be surprised if we declare a victory on abandoning some of the methods and converge into a few methods that would be more suitable.

CHAIR APOSTOLAKIS: That may very well be but I think there is a conflict. I don't think we can resolve this issue now. I'm just telling you that be aware and please don't come back with a final report that praises everybody. It's going to be -- Especially if you send it to us for a quality review. I can tell you what that letter will say right now. So please don't do that.

MS. LOIS: We are very careful with the existing draft report because this is --

CHAIR APOSTOLAKIS: Say it again. Sorry.

MS. LOIS: -- a pilot study.

CHAIR APOSTOLAKIS: The report that you have received is on the pilot, what we've done and --

CHAIR APOSTOLAKIS: No, I'm not criticizing this report, Erasmia. Don't misunderstand me. All I'm saying is two years down the line or whatever when you come up with a final report if you start praising every model and say, "Yes, they considered this, but a little bit here" that's not a worthwhile result.

So let's go on after these happy comments.

MS. LOIS: Okay. I think I covered the motivation for this.

CHAIR APOSTOLAKIS: Okay. Good.

MS. LOIS: So on a very high level what we do here is Halden is performing simulator experiments using real crews responding to transients similar to those models in the PRA and collects true performance data and for this specific study we have two different analysis, steam generator tube rupture and loss of feedwater. Those studies, those runs were performed in November and December of 2006. So we're using the results of those experiments for the whole phase of this study, not just the pilot.

MEMBER ABDEL-KHALIK: At some time during the day, I hope somebody will explain what is so unique about doing the experiments at Halden. Why couldn't these experiments be performed at any plant simulator with the appropriate sort of observations and boundary conditions. And the second thing that I'd like to see sometime during the day is how possible, I understand these procedures have gone through a lot of vetting, how possible errors in the procedures used in these simulators may affect the results of your study.

MS. LOIS: We hope throughout the day some of those questions will be answered and, if not, then at the end of the day we can discuss it a little bit more.

MEMBER ABDEL-KHALIK: Okay. Thank you.

MS. LOIS: But I don't want to --

CHAIR APOSTOLAKIS: Is it at this point where you will address Dana's comment having Swedish crews working on a Norwegian reactor using Japanese procedures. So what does that mean to Americans?

MS. LOIS: One of the presentations --

CHAIR APOSTOLAKIS: You will address that? Fine.

MR. FORESTER: A little bit.

MS. LOIS: And if there are remaining questions, we'll --

CHAIR APOSTOLAKIS: Because I notice you were very careful in the report not to say what the ethnicity of the crews was. Keep going.

MS. LOIS: Okay. This study is three phases. The phase one is the pilot. Beyond that you have the draft report. The SGTR scenario includes two variations, two types of scenarios, one easy one, one more difficult and each one of those analyses has four human actions defined and therefore eight human actions for the SGTR for the pilot which is analyzed to human actions.

CHAIR APOSTOLAKIS: So again, I'm sorry for interrupting so much but this is important. The human failure events were identified by somebody, the team, before the exercises. Right?

MS. LOIS: Exactly.

CHAIR APOSTOLAKIS: Now in --

MS. LOIS: No.

MR. FORESTER: The exercise themselves the actual -- My name is John Forester. The Halden reactor project was already conducting some experiments on HRA/PRA type issues and more from a shaping factors issues. So those scenarios were already developed and, in fact, they were being -- they were running when we were still designing the study. Since there were 14 crews available and some very PRA type scenarios, a complex and a simple, we wanted to capitalize on those for the pilot study. So we wanted to test our methodology. So we worked with something that was already there, but again there were human failure events in there that we thought we could use for the study.

And ideally, of course, we would design

the scenarios ourselves and make sure everything is exactly as we wanted it, but we wanted to test the overall methodology. So those scenarios were already developed.

CHAIR APOSTOLAKIS: But that was for the --

MS. LOIS: But that's not the question. The question is --

CHAIR APOSTOLAKIS: That was for the crews that work on the simulator.

MR. FORESTER: Right.

CHAIR APOSTOLAKIS: But then you also had analysts who used various methods. Correct?

MR. BRAARUD: Yes, it was predefined to that.

MEMBER STETKAR: Will Per's presentation, I notice we have a presentation on the scenarios.

MR. BRAARUD: Right.

MEMBER STETKAR: Will your presentation touch on the definition of the human failure events which is what George is asking about, not -- Everybody knows what a tube rupture scenario is. It's defining the particular human failure event that is the focus of this study which is the salient feature. Will your presentation explain that?

MR. BRAARUD: Yes. It will -- My presentation will explain that and there will some further presentations will also.

MS. LOIS: And also Dr. Vinh Dang's presentation is going to.

MEMBER STETKAR: Because in my mind, that is an absolute key to understanding and interpreting this whole exercise and perhaps the usefulness of it.

CHAIR APOSTOLAKIS: Yes. Okay. So if you guys are going to address it, that's great.

MS. LOIS: Okay. So in terms of status, we've analyzed only to human actions and we created a draft report which is a combined NUREG/IA report and HWR which is an important publication.

CHAIR APOSTOLAKIS: It is not heavy water reactor.

MS. LOIS: I'm sorry.

CHAIR APOSTOLAKIS: It doesn't mean heavy water reactors. You are throwing me off every time I look at it.

MS. LOIS: HWR stands for Halden.

CHAIR APOSTOLAKIS: That's Halden. Okay. I see. Good.

MS. LOIS: Phase two it's the remaining of the SGTR human actions are going to be analyzed and we believe that we're going to have those analyzed by when, by September actually of 2008 and then we're going to deal with the loss of feedwater human actions.

So in terms of input from the Committee on how you define the human actions, I think we will be able to take that into consideration for the loss of feedwater scenario.

CHAIR APOSTOLAKIS: This slide again refers, I mean, is related to my earlier question. I really want to know because I'm pretty sure the Commission wants to know is 2010 the end of this or just of this report. We cannot go on for ten years. I mean, I think somebody will get tired after a while.

MS. LOIS: I believe that 2010 is a good date for the completion of the study as well as part of the studies that will tell us the whole picture about the differences of human reliability methods and how we can converge.

CHAIR APOSTOLAKIS: So 2010 is the time when the staff will send a SECY to the Commission saying, "This is the model we should be using. These are the models." Wait a minute. This is really important. What if the Commission gets tired of this a year down the line and says no more funding? You really have to have the end --

MR. MONNINGER: I think there are two things there. One is the schedule, what's our plan, and the other is resources and funding.

CHAIR APOSTOLAKIS: That's what I'm worried about, John.

MR. MONNINGER: Right. You know the first one was the plan and you asked a very fair question up front, "When are we going to come back to the ACRS? Can we come back to the ACRS in a couple months and give you our entire plan as to how this fits into the broader picture?" So I think we do have to do that, to come back to you and I'm not sure if we can actually say 2010 until we show you the entire plan or not. But that we should present that during the meeting.

CHAIR APOSTOLAKIS: All I'm trying to do, John, here is to sensitive you to the fact that this is not an open-ended project and I don't want you to be surprised if the Commission one day says, "We've had enough of this."

MR. MONNINGER: Right.

CHAIR APOSTOLAKIS: So we have to have a plan. We have to have a target date, what we're going to produce, and I think that's very important for all of us.

MR. MONNINGER: In terms of budgeting for our projects, I mean, for research this is a healthy portion of our HRA budget.

CHAIR APOSTOLAKIS: Yes.

MR. MONNINGER: You know, this individual project. But what we contribute is probably a very small fraction of the worldwide effort to this project. Regardless, we have a long-term relationship with Halden in human performance and fuels and I&C. So this is one of those -- It's partially funded under that.

CHAIR APOSTOLAKIS: I never questioned the fact that you leverage your resources. But I repeat. It's really important for all of us to have a target date, plus I've noticed that as you know the ACRS has several new members and I notice that now a significant number of them do appreciate seeing dates, not just we're going to do this and that. I mean, they are really very pleased when it says, "And this will be completed in June of 2008." So I think it's important for all of us to be sensitive to this, that this is not open-ended, and the Commission has told us many times. This is not a research funding agency. This is a regulatory agency. We should never forget that.

Okay. Enough of this. So where are we?

MS. LOIS: Okay. Overall, we believe that this is a very important study. This is the first time ever that every HRA analyst sits down, analyzes scenarios and it's described in detail how he/she came up with the results and are compared with crew performance data. So, in actuality, it gives us the opportunity to understand how analysts are using their methods. In closure, when you do a PRA, you don't have to give all of the inferences of our method application.

CHAIR APOSTOLAKIS: And we will get more into the details of this.

MS. LOIS: Yes.

CHAIR APOSTOLAKIS: Okay.

MS. LOIS: And then it, of course, gives us the opportunity to perform method-to-data and method-to-method comparison.

CHAIR APOSTOLAKIS: Is EDF participating? Are the French participating?

MS. LOIS: Yes.

CHAIR APOSTOLAKIS: They are?

MS. LOIS: The French are participating.

CHAIR APOSTOLAKIS: Who from EDF? Pierre

LaBot?

MS. LOIS: Yes.

CHAIR APOSTOLAKIS: Okay.

MS. LOIS: So on the positive side if you look at the results of the pilot, we see that the methods and analysts in general are doing a good job and also we note that uncomplicated scenarios are those that probably cannot give us very good insights as to how the methods are applied because you need more challenging scenarios to push the limits of the method and really provide an opportunity to understand how the various methods are applied.

We believe that the pilot produced results that can be used for improving HRA now and Gareth is going to cover that at the end of the day. And we believe that there are some methods that probably will be left out. For example, we don't have many, many teams for each one of the methods to do the variability of the method obligation from team to team and also as we noted before this is at Halden using European crews, etc. So we believe strongly that we have to replicate the study using U.S. crews and Halden is willing to come and do the study here. But we have to have the people to volunteer.

CHAIR APOSTOLAKIS: Replicate the study?

MS. LOIS: Yes.

CHAIR APOSTOLAKIS: We'll never finish. Really? You want to do that?

MS. LOIS: If necessary. Let me put it this way. If we believe that it's necessary.

CHAIR APOSTOLAKIS: Okay.

MS. LOIS: But if you -- If what we come up with at the end of the study is good enough for the Committee and for us, then we'll close the books. We don't have any reason to continue.

CHAIR APOSTOLAKIS: Yes, and again what's good enough depends very much on how this fits in the bigger picture of where we're going to go. Okay.

MS. LOIS: So with that, I will allow Jeff Julius to come in and tell us his perspectives as to why it's very important to participate in the study. Jeff.

CHAIR APOSTOLAKIS: Do we have slides for Jeff?

MS. LOIS: Yes, I have slides for Jeff.

CHAIR APOSTOLAKIS: Okay.

MR. JULIUS: Good morning. Yes, I have about eight slides that just provides an introduction. The first three or four, just let me know when my slides are up.

CHAIR APOSTOLAKIS: They are not now. Now they are.

MR. FORESTER: They're close. There you go.

MR. JULIUS: Okay. Good morning. So this is just an overview. This goes back to your question earlier about where do we stand in the big picture of things. So this is establishing the big picture.

The presentation outline on slide two. You know the Commissioner's charter. So I'm just restating that for whatever, the audience. I really want to talk a little bit about the progress towards the goal and our current plan.

Slide three, part of the importance of the charter is that the importance of HRA is a PRA. I mean, why are we doing this. It's to understand HRA's insights regarding the human error probability and the factors and then which of those factors can I use. Is it a part decision making or is there something I can promote or mitigate in order to manage risk?

The next slide four, the Commissioners' staff requirements memorandum, this is restating this. Work with the staff and external stakeholders to evaluate the different human reliability models in effort to propose either a single model for the agency to use or guidance on which model should be used in specific circumstances.

CHAIR APOSTOLAKIS: This is what I had in mind.

MR. JULIUS: So that's the overall prime directive, if you will.

On slide five, it's the progress towards the goal. So, prior to 2007, the NRC research scientists worked on NUREGs and good practices and the evaluation of method and the HRA database. And then this last year there were two major efforts.

One was the Halden benchmarking first phase which you'll be hearing the discussion on in subsequent presentations. But this first phase focused on the development of the process to collect and compare the empirical data. So we're still -- We're working on tuning our process here. But also in addition to Halden, EPRI and NRC worked on, started on, a joint fire HRA project. So some of the elements of both the prime directive and the lessons learned at Halden are being factored into the joint EPRI/NRC fire HRA project.

The next slide, Slide 6, the integration plan activity, this was our proposed plan. The first step was to establish a team. This was what was presented to this same committee back in, a year ago, March 2007. So I just wanted to present this in order to remind everybody what our initial plan was last year.

The next slide, Slide 7, talks about what's the progress as of February 2008. So the first activity or the first step in the plan was executed to establish a team or a working into this NRC/EPRI Memorandum of Understanding. The second activity, an integrated approach in common terms, it started with both the Halden activities and we just recently completed a workshop on coming to common terms that completed yesterday at 12:00 noon. So the first two activities are underway.

The remaining activities, you're right. We're working on developing the tasks and milestones --

CHAIR APOSTOLAKIS: Wait a minute. You're already on seven.

MR. JULIUS: -- to support the elements of those rules of that plan.

CHAIR APOSTOLAKIS: Can we go back to six?

MR. JULIUS: Go back to six?

CHAIR APOSTOLAKIS: Yes. So under 3(a) what was the answer?

MR. JULIUS: We haven't done 3(a) yet.

CHAIR APOSTOLAKIS: Okay.

MR. JULIUS: So right now, we're writing the plan of doing that. How do we figure that out? Do we survey industry stakeholders as well as NRC stakeholders?

CHAIR APOSTOLAKIS: Okay. Thank you, Jeff.

MR. JULIUS: Okay. So that's the kind of thing you're looking for in this update here in the next meeting or two. Right?

CHAIR APOSTOLAKIS: Yes.

MR. JULIUS: So we're developing a plan because some of that isn't as simple and then while this plan is being developed, we're continuing the development of the Halden benchmarking data and we're starting this NRC fire HRA project and again we're planning to use the lessons learned and this guidance from the prime directive in terms of not developing a completely new method for fire HRA, but how can we adapt or use the best elements of the existing again to promote a common or integrated approach.

So as a summary again, this is just the introductory high level that the Commissioners have chartered the ACRS and the NRC staff to examine HRA methods, propose a single method or guidance on which method should be used in specific circumstances. A project plan is being developed to answer 3 and 3(a) and 3(b) and all the rest of three and four and that includes looking ahead, too. Because you see in the elements of four that to look ahead is to support spatial and external analyses and shutdown and we're using Halden then to promote the establishment of common terms in this integrated approach and at the same time, we're working on a fire HRA. So we have many irons in the fire on, if you will, the HRA side. But we are working to develop this integrated approach.

CHAIR APOSTOLAKIS: Jeff, when you say, "Project plan is being developed and implemented simultaneously," this is a project plan to achieve the Commissioners' goals, to meet the Commissioners' goals? Is that what the project plan is?

MR. JULIUS: Yes. The project plan is basically built up my slide number six.

CHAIR APOSTOLAKIS: Yes. I understand that, but why is the plan still being developed? We don't have a plan.

MR. JULIUS: Because of resource restraints. We have a certain number of analysts and a certain amount of funding and the industry is working on transitioning many plants to NFPA-805 and while we're doing, supporting, other industry initiatives we're also supporting this and there is only -- Like if we had in an ideal world that we could dedicate a team to doing this, we could knock it off quicker than we can given the current situation. This activity is a part-time activity.

CHAIR APOSTOLAKIS: So when are we going to see the project plan?

MS. LOIS: The plan in terms of having a plan, what to do, I think it's ready. Being funded, it's a different story and that's what Jeff is referring to and we can come and brief you any time you want on what we plan to do and probably we'll be able to now that we are at least from the NRC point of view -- Since we don't have the continual resolution of funding issue, probably we'll be able to determine some dates and milestones.

But I believe that components of the plan are being executed as we speak. For example, the first task was what had been proposed to develop a top-down framework as to establish the terminology and that task has been, as a minimum, started with good efforts and Idaho National Laboratory offered a workshop that we conducted this week and we started addressing this issue which is an important aspect.

Now another aspect is the one that Jeff had proposed which is to really identify which regulatory applications are important and sensitive to HRA. That task which we believe is crucial to come up with a good response to the Commission, we haven't -- we didn't start yet and we don't have a schedule yet. But we can establish a schedule.

MR. JULIUS: And the other element of that is this fire HRA really goes to number three and number four. We know that the Human Reliability Analysis during fire is a PRA and HRA application where it is sensitive to the HRA and because of the number, what is it, two-thirds of the industry is doing this transition. Then they need the guidance for how to do that modeling. So that's been a big focus here the start of last summer and will continue through this year.

MEMBER STETKAR: Jeff, this is John Stetkar. I'm one of the new members on the Committee. So I don't have the benefit from some of the history of this whole project.

You seem to make a clear distinction between fire HRA and HRA. Several of your slides talk specifically about fire HRA as if the evaluation of human performance after a fire is different from the evaluation of human performance after a LOCA or a tube rupture or a reactor trip or an earthquake or anything else. Why the distinction as if fire HRA is different from any evaluation of human performance? This again, in terms of the plan, integration of these activities, why is that a separate topic?

MR. JULIUS: Okay. Good morning, John. Welcome to the ACRS. A fundamental approach that we have to the HRA, this identification of the human failure event and the qualitative definition and the quantification, it's the same fundamental approach. What we're seeing in the fire response is that the plants have a wide range of the level of detail in the procedures and the strategy for implementation of the procedures.

So the traditional HRA methods for the level one internal event just following full power operation are primarily based on control room actions and control room actions that follow a procedure-directed response and in the testing and the evaluation of the simulator data is focused on that. What we see in the fire sometimes the strategy is to implement procedures in parallel with the EOPs.

Sometimes these fire procedures replace the EOPs and then when you get to the fire procedures, things that would be a step-by-step procedural guidance in an E-3, for example, for a tube rupture at a Westinghouse plant, if you look at a fire procedure, it may be a one line statement that says locally, manually controlled steam generator level and it's up to the operator then to know the rest. So we're taking a look at how is the plant actually operating and implementing because I mean the goal of the HRA is to reflect as an operated plant and then to see what elements we have in our model to make sure that the model and the approach that we have to identify where it is the same and where it potentially differs and if it does differ, how we would reflect that difference in the HRA method.

MEMBER STETKAR: Okay. Great. Thanks a lot. I appreciate that kind of brief perspective. Thanks.

MEMBER ABDEL-KHALIK: I must say I'm very troubled by the statement that project plan is being developed and implemented simultaneously. Doing things on the fly just doesn't work. You have to have a roadmap of what you're going to do.

MR. JULIUS: I guess let me explain that

further. The elements of the plan, we know what we want to do or what needs to be done. But the level of details in the project plan gets into assigning milestones so that we can see is it going to take to 2010 or is it 2009 or to whenever. So that's a function then of the resources meaning the staff and the funds. That's where I remember that. So we know what the elements are of the plan, but the execution and identification of what milestones and when those are going to be accomplished.

MS. LOIS: However, it is our view that we need to brief the Committee, come down and brief on the whole research plan and give the dates and milestones. So if you bear with us today, we should arrange another meeting where we come in and we provide the details of our activities and how they compliment each other and what are the milestones and dates we have.

MR. JULIUS: Yes, the point of this introductory was just to explain the big picture of where we stand on these different activities. That's the focus of what the rest of the presentation is on and what is being done and accomplished but to let you know we haven't lost sight of these other activities that are going on and we understand that we need to identify the tasks and milestones. We just hadn't prepared to present that today.

CHAIR APOSTOLAKIS: I suspect those words there are not the right word. "Is being developed and implemented" sounds terrible. But probably they do have already something. Anyway, we'll address this some other time.

MR. JULIUS: That might have been my attempt to abbreviate for brevity on the slides.

CHAIR APOSTOLAKIS: Could be. Okay. Are you done, Jeff?

MR. JULIUS: Yes, sir.

CHAIR APOSTOLAKIS: Thank you very much. Who is running the show? Erasmia?

MS. LOIS: The next speaker is Dr. Vinh Dang from Paul Scherrer. Do you want to introduce yourself before?

DR. DANG: I'm Dr. Vinh Dang, Paul Scherrer Institute.

CHAIR APOSTOLAKIS: Who are you and from which school do you come from?

(Laughter.)

DR. DANG: I'm a graduate of the Nuclear Engineering Department at UC-Berkeley and I have my PhD from MIT in Nuclear Engineering.

CHAIR APOSTOLAKIS: Okay.

(Off the record comments.)

II. STUDY METHODOLOGY, ORGANIZATION, PARTICIPANTS

DR. DANG: Okay. We're in the 9:00 a.m. agenda item on the overall methodology, organization and the participants.

In this present I have basically three parts. I'm going to give you an overview of the methodology. Throughout the day you're going to hear many more details about each piece, but this overview is supposed to help you place each of the pieces. Secondly, I will deal with organization and participants just so that you know who actually is responsible behind this work. And, finally, I'll discuss a little bit the work that has been performed in 2007 and just the study phases. I think some of this has been covered Erasmia's presentation. So I'll do that a little quicker.

Overall aims and tasks, well, you know what we're here for. We're trying to assess the HRA methods in light of the simulator data. I think it's important to point out that what we're doing, this comparison using simulator data is something has not been done before, certainly not at this scale. So one of the aims of the work that we've done to-date has been to establish this methodology, to test it and to make sure that the HRA teams that are participating feel that this is fair and a reasonable way of assessing HRA methods.

CHAIR APOSTOLAKIS: So are we going to go into more detail on these things at some point today?

DR. DANG: Yes.

CHAIR APOSTOLAKIS: Okay. All right. For example, I would like to know at some point how you're going to identify the weaknesses. Okay?

MR. FORESTER: We have examples of that.

CHAIR APOSTOLAKIS: Good. And also define scenarios and HFEs comes back to the earlier question.

DR. DANG: Correct.

CHAIR APOSTOLAKIS: Who defines what and when?

DR. DANG: Exactly.

CHAIR APOSTOLAKIS: So all this is later.

DR. DANG: No. Some of that is now.

CHAIR APOSTOLAKIS: Okay. Vinh.

DR. DANG: As far as -- I think this figure here is a bit busy but it gives you an overview.

There are four major parts. One is to define what it is that, here up in the middle, is to define what it is, what performance that the HRA methods are supposed to be analyzing. So that's the scenario HFE and so on.

On the lefthand side are the method applications. So each of the HRA teams are applying their methods based on this information producing, of course, an HEP but also qualitative insights concerning what they expect will be seen and what is driving performance. That's what we call predictive outcomes here at the bottom.

On the right-hand side is the development of the reference data. We have these crews that are participating in the simulator for this scenario. Data is collected and analyzed and eventually it comes down to deciding what it is that you have seen in HRA terms. The fourth part is to make the comparison between the left and the right, predict the outcomes and the experimental outcomes.

CHAIR APOSTOLAKIS: So maybe you should have an oval like the one that says crews on the left to say analysts.

DR. DANG: Yes.

CHAIR APOSTOLAKIS: Is that true? There are two paths?

DR. DANG: Yes.

CHAIR APOSTOLAKIS: Analysts and --

DR. DANG: That is correct. The red path is the analysts' work and the right path is Halden work and the assessment team work in part and the bottom is assessment team work.

CHAIR APOSTOLAKIS: Good.

DR. DANG: Now who is responsible, you see on the lefthand side is the steering group. A number of us are here. You see the Pierre LaBot of EDF is in the steering group and Pekka Pyy of the NEA is also --

CHAIR APOSTOLAKIS: He's not there anymore, is he?

DR. DANG: He's not longer at the NEA. That's correct.

CHAIR APOSTOLAKIS: So what is the steering group doing and what are the other groups doing? I can understand what the other groups on the right are doing. The steering group is doing what?

DR. DANG: The steering group is developing the overall objectives and ensuing that the interest of all the parties are represented in the development of the work.

CHAIR APOSTOLAKIS: Okay.

DR. DANG: The assessment and comparison group, this is where most of the HRA comparison methodology is being developed. The Halden staff is doing the simulator study and the data analyses. You see the crews down here performing in the simulator and on the right there is no space to list the individuals, all the HRA team participants.

MEMBER STETKAR: Vinh.

DR. DANG: Yes.

MEMBER STETKAR: Let me ask. I have to get it in someplace.

DR. DANG: Yes.

MEMBER STETKAR: Who made the decisions for the selection of the particular participants and the methods that we're evaluating? How was that decision process made?

DR. DANG: We put out -- we have been building up to this study over a number of meetings organized by NRC and Halden that took place in as early as 2005 and into 2006 engaging the interests and then as we developed the basic -- as we agreed on the objective of this work we invited people to participate and they indicated that they would.

MEMBER STETKAR: One of the -- The reason I ask and I might as well bring it up now is I noticed in the report, in section two of the report where you talk about kind of the general methodology and things, there are a couple of references to, for example, the PLG SLIM methodology which I --

CHAIR APOSTOLAKIS: I noticed that, too.

MEMBER STETKAR: -- happen to really familiar with and yet that methodology is not one of the methods that has been evaluated here. There are also conclusions that says no human reliability analysis methodology accounts for crew-to-crew variability and I happen to know that that methodology does.

So I was from those perspectives, not necessarily from my own personal familiarity, it seemed curious that there were very few references in section two to specific methodologies. The one that was referenced is not selected for evaluations. So that struck me as being somewhat strange.

MS. LOIS: So, then, from the NRC's point of view of the teams that participated in the method that were evaluated are those that are used mostly in regulatory obligations. Now SLIM although was used tremendously in the IPEs, lately we don't see many obligations of SLIM in the regulatory --

MEMBER STETKAR: I think you'll find that in the United States, but I'll think you'll find that that PLG methodology worldwide is used much more than

CREAM, is used much more worldwide than HEART, is used much more worldwide than some of the other methodologies that are explicitly included in the study. So from the user application perspective, it doesn't seem consistent either.

MS. LOIS: So then actually the way the international community participated in this is Halden asked its signatory members to participate in the study and those that they say yes, they participated and apparently no one that has been SLIM happened to participate. That's why. It was not --

MEMBER STETKAR: Okay. It just seems strange that it was one of the two or three methods that were explicitly references in the document and it was only one that was referenced that was not actually tested.

MR. PARRY: John, let me correct one statement. I don't think this document says that none of the methods takes crew variability into account. In fact, it recognizes that some do.

MEMBER STETKAR: Okay.

MR. PARRY: It says that some clearly do not. But we'll get to that later on.

MEMBER STETKAR: Okay.

CHAIR APOSTOLAKIS: But does the document go beyond that and say that they should, that it's an important -- Or maybe it's too soon to know that.

MR. PARRY: We'll discuss that later.

CHAIR APOSTOLAKIS: But at some point --

MR. PARRY: The conclusions will come out later.

CHAIR APOSTOLAKIS: -- you might say --

MR. PARRY: We lean in that direction.

CHAIR APOSTOLAKIS: Yes.

MR. PARRY: We do. We'll talk about that.

CHAIR APOSTOLAKIS: But that would be an extremely valuable insight by the way from the exercise on the simulator if you come back and say this particular thing that we observe doesn't seem to attract the appropriate amount of attention from some models because then you are helping really people to adopt their models and adjust them and improve them. So that would be very good. But we'll come back to it because we will talk about the timing and all that and that's when all that stuff comes out.

Okay. Great. So really it's a major effort from --

DR. DANG: It is a major effort.

CHAIR APOSTOLAKIS: I mean, you have a lot of people. But these are working part-time, right, as

Jeff said?

MR. MONNINGER: If I could just interject. It is a lot of people and these countries and these organizations are doing that on their own. They are not being funded by the NRC.

CHAIR APOSTOLAKIS: Right.

MR. MONNINGER: And one of the difficulties with these projects is the voluntary nature of it. So though it is a major part of the, maybe not a major part, it's a significant part of the NRC's HRA budget, but the overall project we are definitely benefitting from all these parties. There is no way that we could do something like this on our own.

CHAIR APOSTOLAKIS: Yes. There is no question that this is good. Can you explain to me what is the fourth company from the bottom on the HRA team participants? I've never heard of them.

DR. DANG: NRI.

CHAIR APOSTOLAKIS: What?

DR. DANG: NRI.

MR. MONNINGER: From the bottom.

DR. DANG: From the bottom, Alion Science.

CHAIR APOSTOLAKIS: Who are they?

DR. DANG: John, can you comment?

MR. FORESTER: Yes. They, I think, were formerly MicroAnalysis.

CHAIR APOSTOLAKIS: MicroAnalysis?

MR. FORESTER: MicroAnalysis and Design.

CHAIR APOSTOLAKIS: Where are they?

MR. FORESTER: They're in Colorado and they --

CHAIR APOSTOLAKIS: And they are doing major -- Yes, go ahead.

MR. FORESTER: They participated. They have a software cognitive model that the idea is -- it's used in military applications, things like that. The notion is to model the system and hopefully if you want to change displays in some way or make modifications you can try to estimate how those things would impact the operator's performance and change the timing and the completion. So it's not really a cognitive model, but it's a system model. But it's a software-based tool.

So we got them involved to see if they could take some sort of HRA approach and use that in terms of predicting error probability and also how it would take certain actions to be completed. So it's stepping a little bit beyond HRA methods per se, but began to look at this cognitive modeling approach or modeling.

CHAIR APOSTOLAKIS: Okay. And Politechnic of Milan, right, I see their name for the first time? Who's from the Politechnic?

DR. DANG: This is Professor Enrico Zio and actually these teams down here are in italics because they did not do HRA analyses in this work. Alion Science and Riso and University of Maryland in this context were participating as trying simulation methods to see how simulation methods, I'm talking about computer simulation methods, might be used in the frame of evaluating HRA methods or as an HRA method.

CHAIR APOSTOLAKIS: I see.

DR. DANG: And Politechnic of Milan was working on methods for the data analysis.

MR. FORESTER: Alion Science did try and include a third type model in their simulation.

DR. DANG: Okay. On the next slide --

CHAIR APOSTOLAKIS: Is that the way you're going? You're going to simulation when the HRA -- Well, anyway, keep going.

DR. DANG: I want to do this slide very quickly. I think John has already raised this point. John Monninger has already raised this point which is that a number of the participants are participating based on their own funding and, of course, one of our hopes is that as the results of this pilot get published that other methods that are important might feel that they should be evaluated, that they certainly want to be seen side-by-side with the other methods, and maybe they will then be motivated to come to the table.

CHAIR APOSTOLAKIS: But that would be too late.

MEMBER STETKAR: That would be too late. Essentially you can't -- I mean --

DR. DANG: For the pilot, certainly.

MR. FORESTER: But not for the follow-ons.

DR. DANG: Not for the --

MR. FORESTER: We could include them later.

CHAIR APOSTOLAKIS: See. That's what scares me. The follow-ons will want it repeated with Americans. The Commission will want this thing to finish at some point, guys. We can't talk that way.

MR. FORESTER: Well, we have --

CHAIR APOSTOLAKIS: I mean, it's not research. John, I understand that as a researcher, not just you, but all of us would like to improve, would like do this, would like to do that, but I don't know that the Commission will tolerate that for too long. So let's not talk about follow-ons. Dennis, do you want to say something?

MR. BLEY: Well, yes. George, I think a lot of the reason this was set up the way it is and I wasn't in on the complete organizing part of it is we had a benchmark some years ago that didn't really yield anything that was helpful and there's been a lot of effort here at trying to put this together in a way that we can gain useful information and I think it was felt by most people that getting into it as quickly as possible was necessary and that pilot study was set up without all of the -- was set up not to test everything about the methods and to get a preliminary understanding of will this process really be able to take us to a useful point.

So I think it was wise planning to begin with a manageable piece before you jump into trying to do a much more thorough next step of the study. So I understand there are limits, but if you want it to work well this gives you a chance to correct before you get into the more thorough tests.

CHAIR APOSTOLAKIS: I think it will help us all, you know, maybe at the next meeting if we see that overall plan that there is an end sometime and I don't doubt that you have to try things and so on. That's great as long as we have an ultimate goal and a date.

DR. DANG: May I comment on this as a member of the steering group? This is where the steering group comes in, for example. If there's a development element right now because what we're doing is informative for HRA research. It's informative for practitioners. But the development element is really to develop the comparison methodology and our goal in the study is to assess the methods.

CHAIR APOSTOLAKIS: But that goal is a subsidiary goal to the bigger goal. That's what I'm saying and I guess we are at the disadvantage here because we don't know what you have in mind for the bigger goal. That's all. I'm not going to tell you how to run your business as long as I know that there is an end at some point. That's fine. I mean, you will do it the best way you think is appropriate.

DR. DANG: Okay. This slide I think gives you some details about things that I mentioned when I brought up that diagram slide.

CHAIR APOSTOLAKIS: Do you think or are you sure?

DR. DANG: Four. I'm sure.

CHAIR APOSTOLAKIS: Okay.

(Laughter.)

DR. DANG: This provides with the details on the roles of the four groups study participants. I think the HRA teams, that's quite obvious. They apply the an HRA method or each team applied an HRA method to predict the performance in the PRA scenarios and they generated these predicted outcomes.

The assessment group on the lower left consists of Halden and neutral participants, meaning that were not on any of the HRA teams. We compiled the information package for the HRA teams. So we told them what the scenario was and I'll give you details on that information package in a subsequent slide. There is a question and answer process where we answer their request for clarification of this information package. We assess the testability and the fairness of the predictions that are made so that you don't make a general prediction that is not testable, sort of like a horoscope, and we finally compare the predicted versus the empirical outcomes.

I've mentioned the experimental staff of the Harden reactor project. They actually designed this specific scenario, in particular, how to implement it in the simulator together with the assessment group. They collected the data online and after the sessions and they analyzed this experimental data to come up with at least one level of the empirical outcomes.

CHAIR APOSTOLAKIS: Can you tell us a little bit about the neutral participants? Are these the ultimate decision makers here or who are they? You said that you yourself don't have any model that you are proposing. Is that what you said about models?

DR. DANG: We're neutral in the sense that we are not associated with any of the analyses that are being assessed.

CHAIR APOSTOLAKIS: Any of the analyses?

MR. FORESTER: We're not only HRA team doing the analyses.

DR. DANG: We don't have to defend any of the analyses.

MR. FORESTER: For instance, I did not participate in the ATHEANA analyses.

CHAIR APOSTOLAKIS: I'm sorry. What?

MR. FORESTER: For example, I didn't apply the ATHEANA methodology. Someone else on the other team applied that. I didn't participate in that.

CHAIR APOSTOLAKIS: So you are a neutral participant.

DR. DANG: In that sense. I mean, of course, we're experts on the methods and we have ties to the methods. But when that assessment, sorry, when that HRA team makes a submittal and we judge it we don't have to defend that submittal.

MS. LOIS: And, also, the people that are not connected with methods judged methods. For example, Vihn did not judge the CECA which is the -- And John did not judge the ATHEANA method. So then the ATHEANA analysis was judged by Jeff Julius. Gareth judged --

(Off the record comments.)

MS. LOIS: We tried to have people that are not connected with the methods to evaluate the analysis that comes from the analyst teams.

CHAIR APOSTOLAKIS: Okay.

MS. LOIS: And, Vihn, are you going to cover the October workshop?

DR. DANG: Yes. This is a list of the HRA methods that are represented in the study. I think that in perusing the report, you've probably seen it. I also show who the team members are. So it's more for reference unless you have any questions.

CHAIR APOSTOLAKIS: We don't seem to have that, do we?

DR. DANG: Excuse me.

CHAIR APOSTOLAKIS: We don't have that slide. We don't have it here.

DR. DANG: Did it drop out?

MS. LOIS: I don't know. Probably. That would be my mistake.

DR. DANG: I'm sorry.

CHAIR APOSTOLAKIS: It was not deliberate, right?

MS. LOIS: I don't know. Copying.

CHAIR APOSTOLAKIS: So you have THERP with Bayesian enhancement.

DR. DANG: This is a method being used in Finland.

CHAIR APOSTOLAKIS: Wow. ATHEANA SPAR-H, CBDT, decision trees, MERMOS, PANAME. What is that? PANAME.

DR. DANG: PANAME.

CHAIR APOSTOLAKIS: PANAME. What does it mean?

DR. DANG: A method developed by the IRS and itself which is a regulatory support organization in France.

CHAIR APOSTOLAKIS: HEART, who is using HEART? Yattenthal? Really with all these?

MEMBER STETKAR: Not. I'm familiar with --

MS. LOIS: This is a reference plan.

MEMBER STETKAR: I'm familiar with the ones that are used commercially, not with some of the experimental ones.

CHAIR APOSTOLAKIS: Is anybody using CREAM? Are you testing CREAM?

DR. DANG: CREAM is there.

MEMBER STETKAR: That was my point.

CHAIR APOSTOLAKIS: Are you testing it? Somebody is running CREAM?

DR. DANG: In the study?

CHAIR APOSTOLAKIS: Yes.

DR. DANG: Yes, it's in the right column third one down.

CHAIR APOSTOLAKIS: What's NRI?

DR. DANG: This is the Czech Research Institute. KHRA is a Korean HRA method.

CHAIR APOSTOLAKIS: Okay. Microsaint.

MR. FORESTER: That is Alion software simulator I talked about.

MEMBER STETKAR: The simulations though

were not evaluated as part of this pilot study.

DR. DANG: Correct.

MEMBER STETKAR: Is that because the simulations can't handle performance shaping factors or why if they're part of the scope of the studies?

DR. DANG: At this time, they do not produce HEPs, for example.

MEMBER STETKAR: Okay. That's --

DR. DANG: Okay? So the HRA teams received this information package.

CHAIR APOSTOLAKIS: Where are we now? Is this a new presentation or --

DR. DANG: I'm continuing now.

CHAIR APOSTOLAKIS: Well, we're missing these. How many more do you have?

MS. LOIS: Apologize. How many slides do you have?

DR. DANG: I have nine --

MR. FORESTER: Did you get every other page or something?

MEMBER STETKAR: No, we go the first eight.

DR. DANG: Nine slides.

MS. LOIS: Nine slides.

CHAIR APOSTOLAKIS: Nine more? Yes, we

should have copies of those. Well, maybe this is a good time to take a break while the copies are being made. Can you have the copies made?

MS. LOIS: Sorry about that.

MR. FORESTER: They're right on the back table.

MS. LOIS: It may not --

DR. DANG: I'm following on.

CHAIR APOSTOLAKIS: Okay. This is a different copy. Okay. We all switch to this. Okay. Thank you, John.

MS. LOIS: This includes everything?

MR. FORESTER: Yes.

CHAIR APOSTOLAKIS: It seems like there are two --

MEMBER STETKAR: There's information for the public and then there's information for the Committee.

(Laughter.)

CHAIR APOSTOLAKIS: And he meant more information for the public, not the Committee. Let's make that clear and everything is public by the way. You are on slide --

DR. DANG: I'm on slide ten.

CHAIR APOSTOLAKIS: -- ten. Okay. Very

good. Now we have it.

DR. DANG: Okay.

CHAIR APOSTOLAKIS: Please keep going.

DR. DANG: This information package consists, the first three items are administrative information and general instructions, for example, the agreement forms to respect the privacy of the teams involved not to distribute the procedures which are proprietary and so on. And it gives them an overview of what the pilot study, the structure of the pilot study.

The second item is the essential part for the HRA teams. This is the information on the scenario on the performance environment on the crews and their job aids, meaning the interface, the procedures and so on.

Item four is information on the simulator facility and its characteristics, how operations are performed in that simulator, how manipulations are performed in that simulator, how information is presented, how it is generally used and some of the differences between the simulator and the control room of these crews, performed in the home plant of these crews. Secondly, the scenario descriptions and HFEs which Per will present in his presentation. Third, item number six is the characterization of the crews, their work practice and their training, how they work and seven is the procedures used in the HAMMLAB. So these are like the emergency operating procedures adapted for the HAMMLAB. And, finally, the forms for the submittals, how they should document their answer, their predictions.

I'm going to give you a little bit more detail about these submittals. There is basically three parts to a submittal. The first part is Form A which is open form questionnaire where we ask them to give us the HEP, describe the driving factors and discuss operational expressions. I have a separate slide on the details of Form A. The second part of the submittal is the documentation of an HRA essentially just it would be in a PRA. And the third part is a closed form questionnaire where we ask the HRA teams to present their predictions in a common terminology and for the pilot study, this terminology was based on the HERA taxonomy, the Human Event Repository and Analysis data collection effort which is referenced in NUREG/CR-6903.

CHAIR APOSTOLAKIS: When that -- Go ahead.

MEMBER STETKAR: One thing on -- I looked at the forms briefly. Did -- Let me see if I can put it. Did you ask for any information from the various HRA teams regarding the level of effort that was expended to perform their analyses? Because in a practical sense, I mean we're interested certainly in one area about the resolution of a particular methodology to capture important features of human response and to predict it numerically.

From a PRA perspective, obviously we can't do a multi-month research project for each of 150 operator actions in a PRA. Was information requested from the teams regarding their level of effort?

DR. DANG: Yes, information was requested from the team not specifically on these form.

MEMBER STETKAR: Okay.

DR. DANG: But we asked them to give us a general estimate as to how much they put in. We also know how many team members they had and the time frame in which they did the analysis.

MEMBER STETKAR: Well, if this is part-time, the calendar time can be misleading somewhat.

DR. DANG: This is true.

MEMBER STETKAR: So as long as people are relatively forthcoming in terms of the number of person hours that were implemented, I think that would be useful as an aside if one is to make a decision later to rank order the various methodologies, certainly, two that are ranked equal technically but one is somehow more efficient and would probably be ranked a little bit higher.

CHAIR APOSTOLAKIS: How is -- I'm sorry. Okay. How is HEP, the first one, and in the PRA different? You say documentation of HRA as in the PRA. I mean, when they give you the human error probability, aren't they telling you how that fits in the PRA?

DR. DANG: No. The documentation of the HRA usually includes the HEP.

CHAIR APOSTOLAKIS: They're different?

MEMBER STETKAR: They're just talking about it. They have two forms, A and B.

CHAIR APOSTOLAKIS: But this is Form A.

MEMBER STETKAR: Each HRA team is required to fill in the information in those forms in the format of the form and then as I understand it, the documentation in the HRA is if the HRA team was performing this HRA according to the documentation standards that they apply in a PRA, the narrative, the description.

CHAIR APOSTOLAKIS: And that's not required when you describe the human error probability.

MEMBER STETKAR: It's the same thing.

DR. DANG: No, it would be.

MEMBER STETKAR: It is. It's just a different format of the document.

CHAIR APOSTOLAKIS: It's just a form.

MEMBER STETKAR: It's just a format of the documentation. Form A and Form B, my understanding is to try to provide some measure of consistency among all of the teams so that the assessment groups could compare them side by side.

CHAIR APOSTOLAKIS: Right.

MEMBER STETKAR: Because the documentation in the PRA varies widely from a table with a number in it to a long narrative about what was actually done.

CHAIR APOSTOLAKIS: And the actual simulations really cannot tell you anything about the human error probability. This is just something that you asked the crew, the teams, to do, but it's really separate from the simulation.

DR. DANG: Yes.

CHAIR APOSTOLAKIS: Just want to see what kind of probabilities the various teams come up with.

DR. DANG: Yes, although we do plan to do quantitative comparison at some point.

CHAIR APOSTOLAKIS: Okay. But at this stage --

DR. DANG: -- the simulator data to inform some kind of empirical HEP.

CHAIR APOSTOLAKIS: But at this stage it was just something that you asked them to do. I can see how you can compare the observations at the simulator with the driving factors, right, the PSS and so on? But the actual probability is something that you cannot really say too much about except that if they are widely different you are now in a different domain.

DR. DANG: Exactly.

CHAIR APOSTOLAKIS: John.

MR. FORESTER: I was just going to comment that we can't see what the crews do. I mean we have HEP predictions and when crews do fail then we have that information.

CHAIR APOSTOLAKIS: Yes, but I mean if somebody tells you the probability of this is 10^{-3}, you have no way of --

MR. FORESTER: It has to be a high probability failure to be able to see it.

CHAIR APOSTOLAKIS: Yes.

DR. DANG: So these are the details of the

Form A. Item one is the HEP value as mentioned. In these driving factors or most influencing factors, we asked them to address both positive and negative factors. We asked them to refer to those factors that they identified through the application of the HRA method and not just through their general smartness and using the terminology of the HRA method. So here we asked them to just use the language that they like to use.

And, three, we've asked them to convert the discussion of influencing factors into what really you will see from the simulator or what item is really difficult. So where you might say under item two that the procedural guidance is poor, under three, we would ask them to say what part of the procedure is poor, in what way is it poor, does it mean that the operator has skipped that step or does it mean that they get hung up trying to interpret that step and so on. That's what we mean by an operational expression. What will you see according to that driving factor?

Form B tries to capture the same information but now in a common terminology. That would help us to compare the driving factors across the language. It's sort of a Rosetta stone to convert what one team might refer to as one factor into a common language.

MEMBER STETKAR: Out of curiosity, did you see a high degree of variability in the various teams' abilities to, let's say, fill in these forms or provide the information that you asked for in the format that you asked for it?

DR. DANG: I think --

MR. FORESTER: We did see some variability to the point where we thought it was worthwhile to summarize the information they had provided.

MEMBER STETKAR: I noticed that. Okay.

MR. FORESTER: But overall it was pretty good and most of what's in the summary actually came from what the teams provided. But we tried to make sure that there's a fair representation of what --

MEMBER STETKAR: I notice that you've gone back to the teams to ask them whether your interpretations are what they were really trying to say also. Right?

MR. FORESTER: That's correct.

DR. DANG: Okay. Now I'm just going to review a little bit what we did in 2007. So our aims were to establish this methodology and to test it, to obtain both some initial results concerning the HRA methods as well as feedback and improvement of the study methodology itself.

In terms of tasks, the collection of data as far as up here on the upper left on slide 15, we prepared an information package for the HRA teams in January of 2007. This was given to the teams and then it went through this question-and-answer clarification process, do they understand what we said about the scenario. On the left began the data analysis on the data that was collected and this feeds into this comparison.

Now what's very important to point out is that in preparing the information package which was done in January none of us who prepared the information package knew what had happened in the simulator to avoid any kind of bias in describing the scenario based on knowing what happened in the scenario. In fact, we did not learn, those of us who were assessing, about the performance until after we had already summarized the HRA submissions and exchanged clarifications with the teams about what they were predicting.

MEMBER STETKAR: In the information package and reading through some of the HRA team analyses, it sounds like -- Is someone else, by the way, going to describe more what was in the information package or is this the opportunity to ask about that?

DR. DANG: I think this is the main opportunity.

MR. FORESTER: We will talk a little bit more about the summaries that were made.

MEMBER STETKAR: Okay. No, this is the input to the HRA teams. Did you provide in the information package time lines, traces of the actual plant perimeters as a function of time through the scenario?

DR. DANG: Yes, we did.

MR. FORESTER: Or a sample.

MEMBER STETKAR: Or alarm readouts, things like that?

DR. DANG: Yes.

MEMBER STETKAR: Were those --

MS. COOPER: Can I comment on that?

MEMBER STETKAR: -- traces, let me finish, Susan.

MS. COOPER: Dennis will want to say something, too.

MEMBER STETKAR: Yes, I'm sure you will. Let me just finish.

MS. COOPER: Okay.

MEMBER STETKAR: Were those traces from an unmodified scenario or were they results from operators interacting with the simulator?

DR. DANG: They were the result of Halden staff interacting with the simulators. So we did not --

MEMBER STETKAR: So they were --

DR. DANG: There was not an unmodified scenario.

MEMBER STETKAR: So they were biased scenarios.

MR. FORESTER: They were the actual scenarios that the crews saw.

MEMBER STETKAR: Well, but they had human interaction in them.

MR. FORESTER: They had human interactions. Right.

MEMBER STETKAR: So that affects on pressures, temperatures, levels presented to the HRA teams were not the plant performance unmodified by human input.

DR. DANG: Correct. That's right.

MEMBER STETKAR: So why is that a blind set of input for human reliability analysis because it is not actual plant response that the operator would see. It is a plant response modified by operator interaction.

DR. DANG: If you did not interact with the simulator, then for the later HFEs the situation that you enter in unmodified is not the scenario that you --

MEMBER STETKAR: Okay. But that's a different test. This is a test for the first HFE.

DR. DANG: I am not sure what interactions were done up to the first HFE because the traces cover all the HFEs.

MEMBER STETKAR: I understand that and as people interact with the simulator the plant performance on the simulator changes and therefore the input to the HRA team trying to understand the scenario progression is different from what the operator would actually see in the real plant environment. That's my whole point is that if you're presenting the HRA teams with the development of an event scenario which is plant response and asking them to evaluate the performance of the operators within that environment that environment should not be perturbed by some other operator performance I would think unless you're very, very, very careful to be absolutely sure that the entire scenario progression up to the point where you're asking the teams to evaluate this particular response has not been perturbed and what effort was made to do that.

MR. PARRY: Actually, John, I think you're right in the sense that I think, at least, one of the teams used that trace as in-going information.

MEMBER STETKAR: I mean, that's typically what many people do use is input information.

MR. PARRY: No, but specifically though the response time.

MR. FORESTER: Right, but I think there are few actions and I'll defer to the Halden guys, but there are few actions like in some cases tripping the plant or the basic response actions associated with a trip and I guess in the case of the complex scenario there's a main steam line break. But those kinds of things were responded to in where we got the traces. But then those are sort of basic actions and there's an assumption those will be done and then, given that, we looked at what the steam generators were doing over time and that was what they had to respond to because that was the main event.

MEMBER STETKAR: It's interesting that you mention tripping the plant because reading through the team responses at least two of them seemed to focus on the time to trip the plant. At least two and perhaps several others focused on the time to trip the plant as a key time in their time-based analysis and if the traces had information about how long the operators took to trip the plant, that seems to be an input bias.

MS. COOPER: I agree, John, and that was going to be my point. The ATHEANA team was one of those that used that and the time at which the plant was tripped was critical because it took up time, other time, for doing other steps.

MEMBER STETKAR: That -- I read it in ATHEANA, Susan, but I also read it in a couple of the other methods --

MS. COOPER: Yes, there were other people, too.

MEMBER STETKAR: -- that looked at time reliability correlations that keyed on manual trip of the plant as the key cognitive response to recognize there was a problem. So I think we need to probably move on here, but my point is that the information provided to the HRA teams is obviously really critical and if the experiment did not carefully think about the format and the form of that information it may, indeed, effect some of the results. So I don't know whether it did or didn't but that's why I asked the question about what was in there and was it a pure hands-off type scenario trace of plant parameters.

MEMBER ABDEL-KHALIK: But how useful would the information be --

MEMBER STETKAR: Very useful.

MEMBER ABDEL-KHALIK: -- if it didn't include any operator actions in a sense that -- You know, I can understand your point.

MEMBER STETKAR: If I'm in the control room and they're asking me to identify that I have a tube rupture and isolate the steam generator, I want to know what the control room's parameter displays are telling me.

MEMBER ABDEL-KHALIK: Right. Up to a certain point when the conditions are modified significantly by subsequent operator actions.

MR. BRAARUD: Can I comment?

CHAIR APOSTOLAKIS: Absolutely. Yes. Please.

MR. BRAARUD: Per Olvind Braarud, Halden. This was actually quite a difficult issue that was discussed quite a lot before we made the examples from the interface and we also found out that we needed to include the human response system to make the progress as realistic -- I mean, the information presented later in the scenario would be something relevant to what the crew who had to predict would actually see in the simulator. But it was discussed quite a lot what kind of model of this operator should we then use because it will give an example. And the solution was that we had thought that this is a typical crew based on the experiences from the instructors. This is how long they are typically experienced that the crew used in the first part of the diagnosis procedure, for example. So if we want to include human actions we have to choose a model.

MEMBER STETKAR: My only point is that when you're performing a human reliability analysis for PRA and you're evaluating in this case theoretically the first accident in a developing -- the first action in a developing scenario, the first operator action in a developing scenario, as this action is presented in the study, then that action ought not to be perturbed by preceding operator actions, success or failure of preceding operator actions, especially if you're asking human reliability analysts who use a wide variety of different methodologies.

Some keyed on time. Some keyed on plant

parameters. Some keyed on procedure steps to evaluate the likelihood of success or failure of that first action, even by perturbing the scenario by allowing the operators to manually trip the reactor and I don't care whether they're average operators or good operators or bad operators or average times or abnormal times, conservative times, whatever that means.

You've already perturbed the scenario. You've perturbed the input information that some teams may use and I don't know what teams may use what information in their analysis. But some teams may use that information.

MR. FORESTER: There are two mentions, two sides of that. I think you're making a very good point and we do discuss some problems with the time criterion that we used and when they trip the plant and those kinds of things were issues and that's a lesson learned essentially from the pilot study. We do need to address that.

But I do think in order to be able to do this research that you do have to assume that there will be some basic actions taken before you get the human failure event that you're focusing and the teams that are doing the predictions have to understand and they have to take into account what they think the crews will actually do. We tried to give them some information with respect to that based on past performance, you know, how long would it take before they trip the plant, about how long does it take them to get through a certain set of procedures.

So we try to give them information about it, so I have an expectation for what the crews will do. But there is an assumption that the crews will do these sort of basic actions before they get to it and they have to take that into account in terms of when they think the crews will do those actions or whether they will do those actions and that's part of their prediction.

And the way the scenarios run, well, I think that's all I need to say on that. I think this part will work, but we have to be very careful about how we define these in the future.

MEMBER STETKAR: I understand if you're evaluating, for example, the second action in your string that if now I'm evaluating the condition of likelihood of success for the second action, start the cooldown, for example, I think it's reasonable to say that here is the progression up to that point. The operators, they trip the plant at this time. They started to reduce feedwater and so forth and now evaluate the condition of likelihood of performing the next step. However, in this particular case, it's the first operator action.

MR. FORESTER: Well, it's the first operator action we're focusing on.

MEMBER STETKAR: The first operator.

MR. FORESTER: The first operator action that is taken.

MEMBER STETKAR: But some of the teams focused on actions that were already completed successfully within the plant response that you provided to the HRA teams, namely, tripping the reactor.

MR. FORESTER: They had to make an assumption about when they would trip the plant. That was part of their analysis and that could impact how they judge the probability of success on following the actions. So that was part of their analysis and we need to be able to give them if we're going to have that kind of time criterion set up as we did we need to be able to give them reliable information that doesn't bias what they do later but some way to estimate what the crews will do.

MEMBER STETKAR: We should probably go on because of time here.

MEMBER ABDEL-KHALIK: Wasn't there a technical basis for the EOPs?

MR. FORESTER: The EOPs are based on the Westinghouse standard guideline.

MEMBER ABDEL-KHALIK: Right, but why wasn't that provided as the input instead of actual operator response?

MR. FORESTER: We gave them procedures. They had the --

MEMBER ABDEL-KHALIK: Not just the procedures, but you know the rationale and the timing for different actions within the procedures.

MR. PARRY: I don't think that you can extract that from the basis for the procedures. I mean, they are success based. They are -- At least, I don't believe so. Right?

MR. BLEY: It's not that they're success-based. Dennis Bley. It's that when those were developed we were actually in another organization I was with. We were working with Westinghouse on the PRA at the time much of that was developed. They made lots and lots of thermohydraulic runs. They made pieces of PRA to address some of those points and the richness of all of that information was really briefly summarized in the basis documents. So they don't really provide you this.

But if I might just take a second to come back to what John said. Your points are right. But if you're doing a PRA, usually you get the plant to run some cases for you and you watch what they do. Just an aside, it turns out that the timing for some of those key things you're focused on, if you look at what the real operators did, were more misleading than helpful. But they are a source of bias.

Also the way that information came to the teams was in a package with a short time to ask questions about it. Responses to questions came back to us. But at least for some of the things for our group, you almost had to divine what people were doing where because that was clearly marked as operator to this hearing here. You saw alarms and things and you could infer.

MEMBER STETKAR: So the information from the simulator didn't have a little mark on it saying operator tripped the plant, operator --

MR. BLEY: But you did see a trip and possibly for that one you saw it was an indication it was a manual trip and the data. It was a computer printout of all the alarms and the time were there.

MEMBER STETKAR: They were there? Okay.

MR. BLEY: The running clock was there. So it was embedded in the --

MR. FORESTER: But that was just one example.

MEMBER STETKAR: Well, I just -- You happened to mention reactor trip and I picked up on that because I noticed that two or three of the teams seemed to use that as a key action in their evaluation of the cognitive response part of the process based on possible biases.

MR. BLEY: In an HFE based on time something that can happen over a -- of time.

MEMBER STETKAR: Right.

MR. FORESTER: And that certainly is important. That certainly led some teams to make different assumptions that did affect the results.

CHAIR APOSTOLAKIS: Now, Vinh, can you wrap it up as quickly as you can?

DR. DANG: Yes. Okay. One last comment about the information basis, over here on the right in this clarification process, it's important to note that every team saw every answer, every question and every answer, from the other teams. So that means that it's not a one-to-one assessment group or a study group to one team. But if one team asks a question, all the teams saw the question and the answer.

Okay. So we performed this comparison. The arrow at the bottom goes to the next slide where we reported on this in the draft report in the version from October of 2007 that describes the study methodology. This report is very similar to the report that you have received for this briefing.

Following the release of that draft report, the U.S. NRC hosted a workshop here in Maryland where the assessment group, Halden and all the HRA teams, were represented and now we're in the process of finalizing that report based on the workshop results.

That workshop discussed the results concerning the HI methods and including both what we did in terms of the data analysis and the comparison methodology. It also went into the lessons learned so far from the pilot of the methodology in terms of the comparison for the assessment group, for the experimental group in terms of data analyses in the context of an HRA study and bring in what the HRA analyst felt could be better in the information package and so on and there was feedback and discussion.

And then we decided on what should be done with the rest of the data. And you've seen the report. I'll skip this slide. And this slide you've also seen presented by Erasmia Lois, the study phases. So I think I can just jump to the last --

MS. LOIS: I do want to make a comment. The report that you received is a draft report and we plan to publish it, to make it available in ADAMS to the public. So then we would like to have your feedback to revise the report as needed. But we're still revising it.

CHAIR APOSTOLAKIS: You're asking for a letter?

MS. LOIS: No. Just the discussion.

(Laughter.)

MS. LOIS: On the basis of today's discussion, we're going to potentially do some revisions.

DR. DANG: Okay. My last slide is the overview of the rest of this briefing. We've done this introduction with the overall study methodology. Next you will have the scenarios, followed by the summaries of the NRA analyses, what we done with the submittals and predictions, then a discussion on simulator data, a data analysis that we developed in the reference data. Obviously the reference data, the reference scenario, here on the left is going to both these tasks, but this is the order of the presentation. Following the simulator data analysis, we have the comparison results and finally we wrap up with insights, ongoing work and the next steps.

CHAIR APOSTOLAKIS: So all this will happen today?

DR. DANG: Yes.

MR. FORESTER: That's the plan.

DR. DANG: That's the plan.

CHAIR APOSTOLAKIS: Okay. Are you done?

DR. DANG: Yes, I'm done.

CHAIR APOSTOLAKIS: Thank you very much and we will take a break down and come back at 10:30 a.m. Off the record.

(Whereupon, at 10:14 a.m., the above-entitled matter recessed and reconvened at 10:33 a.m.)

CHAIR APOSTOLAKIS: Okay, we are back in session and the next presenter is -- I can't see. Nobody is helping me.

MS LOIS: Can we move the --

MR. BLEY: Your name stays with you.

CHAIR APOSTOLAKIS: Mr. Per Braarud, okay, from the Halden Project.

III. SIMULATOR, SCENARIOS, CREWS, HUMAN FAILURE EVENTS

MR. BRAARUD: My name is Per Braarud from the Halden Project. My background is simulator studies, human factors, control room and I'm going to speak a bit about the conditions for the data collecting scenarios, the participants and so forth.

Very quick outline about overview of presentation. A short menu from the lab, say something about the crews, their training, scenario overviews, something about the HFEs and a little bit about the procedure steps for the SGTR scenario.

So the process simulated, it's per core director, simulates **** (10:34) from late `70s.

MEMBER STETKAR: Prelude plant?

MR. BRAARUD: It is a Prelude Plant. The lab used a computerized interface and the procedures used, they are from the participant's home plant based on the Westinghouse procedure package but they are adapted through the computerized interface, how to find the information and they are also adapted through the process simulator when it comes to parameters and criteria.

MEMBER STETKAR: Participant's home plant, you're going to talk about the crews a little bit more, so I'll ask you then.

MR. BRAARUD: Yeah. So this shows a picture of the area inside the control room or the laboratory. It has a large display and it has -- for this study, it has three operator work stations. The operators can select formats, do the control of components and objects.

MEMBER ABDEL-KHALIK: Big picture, how does the layout of the control room effect the results of this study?

MR. BRAARUD: Important question but also quite big question. I will come back to that later but the comments from the participating crews is that they thought this simulation was realistic. It was quite similar to operating the scenarios as they would have them in their training simulators.

MEMBER ABDEL-KHALIK: But if the operator has to walk to a given, you know, location on the boards to take specific action, this is not representative of a lot of U.S. plants and does that effect the results?

MR. BRAARUD: That is correct, but the purpose of this simulation is to have reference data to compare the HRA analysis against and the conditions both the interface, the control room, the way you are working, the procedures, where they're described in the information package. There is also given --

CHAIR APOSTOLAKIS: Yeah, but do they have to do extra steps, right, that's what you mean, sir?

MR. BRAARUD: Right.

CHAIR APOSTOLAKIS: If they have to do something extra like go behind the console, that probably --

DR. LOIS: Yes, here we're testing the methods and not the applicability of a specific method in a specific plant.

MEMBER ABDEL-KHALIK: But conceptually, you're trying to make this as realistic as possible in order for the data to be as representative as possible. And it would seem to me that the layout of the control room has a major impact on how the operators will respond.

MR. FORESTER: But they're asked to predict performance in this control room. So that's what the HRA teams are predicting and they're giving up to like five hours of training on interacting and performing actions in this control room so they are trained on it.

MEMBER ABDEL-KHALIK: I fully understand that.

MR. FORESTER: But the HRA teams, again, they're predicting performance here. They're not predicting performance in their home plant.

MEMBER ABDEL-KHALIK: Ultimately, what we would like to know is how the operators will respond in a real control room, not in the Halden simulator. This doesn't tell us that.

MR. BRAARUD: Okay, the study doesn't focus on that directly, but we have all the projects. We have the focus on comparing research from the lab to studies in other training simulators in other operational places.

MEMBER STETKAR: Were the HRA teams provided information about the layout of the simulator and, you know, this type of information?

MR. BRAARUD: Yeah.

MEMBER STETKAR: So the HRA teams understood that within the context of the exercise, this is the control room.

MR. BLEY: And we had a -- you sent us a movie that we could watch.

MR. BRAARUD: Yes, I'm coming to it.

MR. FORESTER: And the crews had stated that they didn't see this as really effecting their performance all that much for whatever that's worth.

But again, that wasn't the focus of it. I mean, they had to predict the performance here not in the --

MEMBER STETKAR: The important thing is if the crews were operating in this environment and the HRA teams clearly understood this environment. That's --

MR. BRAARUD: That's the key for this purpose but I understand impression is very important and we are investigating that also in other forms.

CHAIR APOSTOLAKIS: Okay, let's keep going. We're kind of behind.

MR. BRAARUD: Okay, I would like to show you a excerpt from a video. This video was given to all the HRA teams to explain or illustrate the operation in the control room. This video shows a quick explanation --

CHAIR APOSTOLAKIS: How long is it?

MR. BRAARUD: Six minutes. Should I take examples from it very quickly?

CHAIR APOSTOLAKIS: Can you shorten it or -- don't skip it. Do you want to skip it?

MR. BLEY: Why don't you watch it till you decide --

MR. BRAARUD: You can decide to stop it when you --

CHAIR APOSTOLAKIS: We can always stop it, yeah, okay.

MR. PARRY: When you fall asleep, we'll stop it.

CHAIR APOSTOLAKIS: Good, good.

MR. BRAARUD: The video quickly explains the control room which is directly into a demonstration scenario where there's spurious actuation or safety injection, just to illustrate how it works and the crew is Halden staff. It is not the licensed crews.

(The video was played.)

CHAIR APOSTOLAKIS: Yeah, I got the flavor. As far as I'm concerned, you can stop. Anybody else wants to continue?

MR. BLEY: As I recall, they did have -- they walked through a bit of a scenario just to show how the communications there works and how they can actually operate the equipment.

CHAIR APOSTOLAKIS: I see. And you're leaving all this with us?

MR. BRAARUD: You can have a copy of it, yes.

CHAIR APOSTOLAKIS: We may watch it another time.

(Off the record comments)

MR. BRAARUD: So this gave a short illustration without the song but this is how the crew team operates in the control room. It's quite comparable to the conventional control room in terms of communication, the roles, how they use the procedures, and so on.

CHAIR APOSTOLAKIS: Okay.

MR. BRAARUD: Quickly about the simulation setup. This is also related to what is the difference between this simulator and a training simulator. We put more weight on having a run schedule that we follow for all crews, so that the conditions are similar for all crews.

CHAIR APOSTOLAKIS: What schedule?

MR. BRAARUD: A run plan for the scenario. For example, the failures are implemented at the same process status or at the same timing criteria for all runs. And some scheduled or what they call a feelibrator (phonetic) for example, that they role play. They have some similar time for the feedback. So that's all conditions for all crews are similar.

CHAIR APOSTOLAKIS: Very good.

MR. BRAARUD: That's my point. So the purpose is not the training. We try to run it as realistic as possible for all crews, similar as possible. So --

MEMBER ABDEL-KHALIK: How are those times determined? If a step in the procedure calls for the control room operator to call an aux operator to do some action in the aux building --

MR. BRAARUD: That timing is estimated by -- from experience of instructors, so it's an approximate time. For example, if a field operator is ordered to check auxiliary feedwater valve pump for example, it is an approximate time estimated from their experience. So it's not a very detailed analysis.

MR. PARRY: Per, I think it's the time of response that's being controlled, not the time when the action is requested.

MR. BRAARUD: Yeah, it's the response time.

MR. PARRY: Yeah, it's the response time that's being controlled.

MEMBER STETKAR: When you run these simulations, do you also introduce distractions that would occur in the control room, for example, phones ringing, people coming in to ask questions and things like that or are the crews left alone?

MR. BRAARUD: That can vary but in this study we didn't have any extra disruptions other than those the crew initiated themselves. So the staff actually role-played those requests that the crew called for. Also the simulator has a good equipment when it comes to logging all process parameters, all the actions that the crew performs through the interface for a good recording of what is going on for each crew.

We also record audio/video, also quite common for training simulators. The participating crews in this study was licensed operators from a Swedish PWR. Consisted of 12 or 14 crews from two different units. The study consisted of a supervisor, reactor operator and assisting reactor operator. The normal configuration is including a balance-of-plant or turbine operator which did not participate in this study.

MEMBER STETKAR: I'm not familiar with the plants in Sweden. All 14 crews did -- they came from the same plant?

MR. BRAARUD: Same site, two different units.

MEMBER STETKAR: Same site, two different -- okay. Are these units also three-loop plants.

MR. BRAARUD: They are also three-loop plants.

MEMBER ABDEL-KHALIK: But this is not the normal configuration that you -- that they man their control rooms.

MR. BRAARUD: They have one more operator, a turbine or balance-of-plant, I guess, it's often called here.

MEMBER STETKAR: And that person, you know, is very busy during the steam generator tube rupture event.

MR. BRAARUD: Yeah, we -- the turbine operator will be occupied with the turbine site and the important tasks for the -- let's say the reactor side or the important steps in the procedures for SGTR is actually performed by these three person. So it was about what persons were available for the study. If we had a turbine operator available, we would, of course, include him.

MEMBER STETKAR: So although we have 14 crews, we actually only have one common organization or common training, common let's say philosophy for implementing procedures so that this is not really a broad input from a variety of different crews. It's different flavors, if you will, of one organization.

MR. BRAARUD: Yeah, that is good. They have the same training, same background.

MEMBER STETKAR: One other question that I had on differences in the plant. This is quick. It's apparently, I think, from reading through the scenarios that the Halden simulator has N16 monitors and it sounds like they're probably downstream from the MSIVs. Is that the same with the operating plants, the plant that the operating crews came from?

MR. BRAARUD: I can't tell you that for sure but there were not comments from the crews that this was unfamiliar to them. So I will expect --

CHAIR APOSTOLAKIS: Do we have assistant reactor operators in the United States? What do we call them? I'm familiar with senior reactor --

MEMBER STETKAR: It depends on the plant. It depends on the plant manning.

DR. GORDON: No, but wait, do we call them that?

MR. BLEY: Sometimes, they're both reactor operators.

CHAIR APOSTOLAKIS: Yeah, that's what --

MEMBER STETKAR: You can have two reactor operators. You can have a shift supervisor and a shift foreman who can help out. It depends on --

CHAIR APOSTOLAKIS: So you can have two reactor operators. One plays the role of the assistant.

MEMBER ABDEL-KHALIK: Sure, this is the primary side and the secondary side.

CHAIR APOSTOLAKIS: Okay, all right.

MEMBER ABDEL-KHALIK: The point that John raised about, you know, all these crews come from the same utility. You have the same training. They have the same interpretation of the procedures. They have the same sort of reward system, jumping steps in the procedures, may not really be the same profession that most utilities would have.

MR. BLEY: But if we were doing a PRA -- I mean, an HRA for a particular plant, all the people we'd look at would be from that facility. So it's -- for this kind of an exercise that they all have that same background is probably appropriate. But it might not be typical of what we would see here in the United States.

MEMBER STETKAR: What might not be typical is the evaluation -- the HRA team's evaluation of expected operator performance because each HRA team is probably trying to think about either an international generically crew, if something like that exists, or a crew performance within the context of their own national experience. You have HRA teams from various countries. Perhaps the philosophy of implementing the procedures in the United States might be different from the philosophy of implementing the same procedures in Switzerland or in Britain or in the Czech Republic or in France, given the same procedures.

So that the ability of HRA teams from those various countries to predict the performance of this -- of these particular Swedish operators, even perhaps using the same procedures could be a source of variability.

CHAIR APOSTOLAKIS: I think there are a lot of issues that are being raised by the members and clearly these guys cannot test everything.

MEMBER STETKAR: No, that's true.

CHAIR APOSTOLAKIS: So the best way to approach this is to state clearly up front what you did and what was left out, what was not tested.

MEMBER STETKAR: Well, the only question was that I didn't recognize reading the report, that all 14 crews came from the same site.

CHAIR APOSTOLAKIS: That's true, as I said --

MEMBER STETKAR: Fourteen crews selected at random from 14 different countries, all of which having PWRs, is a different experiment than 14 crews from one site.

CHAIR APOSTOLAKIS: Speaking of this interpretation, I remember there is a statement somewhere in the report, correct me if it's not a good recollection, that sometimes the crews did something other than what the procedure said and they were proud of it.

MEMBER ABDEL-KHALIK: Praised for it.

CHAIR APOSTOLAKIS: Praised or proud?

MEMBER ABDEL-KHALIK: Praised, I think is the term that was used.

CHAIR APOSTOLAKIS: Praised. Is that --

MR. BRAARUD: Was that a --

CHAIR APOSTOLAKIS: I don't remember.

MEMBER ABDEL-KHALIK: I remember reading it somewhere, too.

CHAIR APOSTOLAKIS: Would an American crew be praised if they did something other than what the procedure says?

MEMBER ABDEL-KHALIK: I think they would be pilloried.

MEMBER STETKAR: Well, not necessarily.

It depends on where you are in the procedures. The procedures are not important.

CHAIR APOSTOLAKIS: So it depends on the action, you mean, what they actually did.

MEMBER STETKAR: Right, that's right.

CHAIR APOSTOLAKIS: I think we tend to be more compliance oriented here. They are expected to follow the procedures, is that true, as opposed say to the French?

MEMBER STETKAR: To the point at which the procedures comply with their interpretation of what's happening in the plant.

CHAIR APOSTOLAKIS: What's going on, yeah.

MEMBER STETKAR: The procedures are never meant to be 100 percent complete for every possible scenario. So you know, there must be some allowance for you know, independent thinking.

CHAIR APOSTOLAKIS: Thinking.

MEMBER ABDEL-KHALIK: But there are --

MEMBER STETKAR: But generally, I mean, if you depart from a procedure or you skip steps in a procedure because you believe that you know what is happening, you're generally penalized for doing that because the procedures are, in fact, designed to methodically lead you through a thought process.

MR. BRAARUD: But also this is a Swedish crew, they are expected to follow the procedures, so that is the basic assumption. So if they are going to deviate from the procedures, they are going to have a good reason and it's typically discussed in the crew and approved by the supervisor, for example, to want to deviate in the procedure or skip a step in the procedure and so on and so forth. That is not at all a normal phenomena.

CHAIR APOSTOLAKIS: Okay.

MR. MONNINGER: If I could just throw in a comment or two, this is John Monniger. I think with regard to the strengths of this experiment, this approach versus, you know, it's direct applicability, you know, to U.S. France, I think one thing that we have to keep in mind is, you know, this actual run was already planned prior to the time that this experiment came along and that you know, conducting this test, we would not be able to do this test or this experiment.

It's really an opportunistic opportunity to take advantage of this. So there is -- you know, I think we have to -- you know, whereas we would like the ideal approach, I think we have to recognize that there are some strengths in it and there are some things that maybe aren't exactly what we would like, but given our -- you know, the resources and bringing together these people, I think there's --

CHAIR APOSTOLAKIS: No, there's no question about it as long as you state it clearly.

MEMBER STETKAR: My only point was -- and I agree with you. I think it's a wonderful opportunity to use resources that are available in a context to provide some really, really useful input. My only point is that when the evaluations of the various methods, the strengths and the weaknesses of the various methods are made, that the evaluators, the assessment teams, should be sensitive to the fact that although this is an experiment, it is a well-defined, well-controlled experiment both in terms of the scenario of the definition of the human failure events and the inputs from these teams.

This is a controlled set of teams, and that fact, because the teams are controlled, that may lead to wider variability in the HRA or less variability within the team, let's say.

MR. FORESTER: Yeah, the teams may bring their own bias.

MEMBER STETKAR: That's right and they probably do.

MR. FORESTER: On the other hand, HRA

methods should be able to predict performance for any country.

MEMBER STETKAR: If the HRA teams knew that these were Swedish operators and were familiar with the kind of Swedish training philosophy and --

MR. FORESTER: We tried to give them all that information. We tried to hold that constant. Since we gave them information about these crews and how they performed in their plant, so we tried to give them -- created more of a baseline in the sense that it was constant for everyone.

MR. BLEY: May I read the sentence, at least from the stuff we got that is applicable to what John asked?

MR. FORESTER: Sure.

MR. BLEY: It says, "When it comes to taking some actions ahead of time, there are no clear directions as to whether this is allowed but we believe that such actions would be praised at training if it made the situation better."

MEMBER STETKAR: "We" being whom?

MR. BLEY: The "we" is the Halden group who was informing us about how the operators carry out their actions in the plant.

MEMBER STETKAR: So your sense is that the operator is -- the Swedish operator, these crews anyway are rewarded for doing things rapidly if they're right.

MR. BRAARUD: I think also a very important lesson, we have to look at the wording of the information where we will send it out today for 18th because the wording of this sounds like it may follow and proceed less strict than they actually do. So it's important that this information package is competent and the formulations are --

MEMBER STETKAR: I think that is really important because I went through license training 25 years ago and in those days before the days of symptom-oriented procedures, before a better way of thinking about operator interaction, I can guarantee you, we were rewarded for making rapid decisions very -- correct rapid decisions. That was a primary reward system in our training. That to quickly draw a conclusion and to react to that conclusion was a very important reward system. That is now contrary to the basic philosophy, at least in many countries of the way that the procedures are designed and implemented, that people are not supposed to be rewarded for making quick decisions. How much that reward system effected these particular crews' responses, I have no way of measuring that.

You may have more insights on that, but it is something -- I'll just reiterate, it is something that when the assessment teams evaluate the variability in the different HRA predictions, that may be a factor, that just may be a factor.

CHAIR APOSTOLAKIS: All right, next.

MR. BRAARUD: Briefly, overview of the experience of the different positions. This is only to show that we have typically a broad sample of what you would expect from a crew, some where new operators, some were experienced in all positions.

MEMBER ABDEL-KHALIK: But these supervisors were previously reactor operators so that the person with one year of experience as a supervisor has a lot more experience as an operator.

MR. BRAARUD: Yeah, this is the actual experience in the given position when they started as a supervisor. This here.

CHAIR APOSTOLAKIS: So there is an assistant reactor operator with 25 years of experience. The poor guy was never promoted or --

MEMBER STETKAR: Plants in Europe are very stable.

MR. BRAARUD: One part of what the crew

goes through before they run the scenarios is actually training in the laboratory and this training is aimed at explaining how the interface works and if there are some differences between the simulated plant and their home plant and they focus then on things that are of importance for the scenarios to be run. And it's typically like they have a more theoretical goal and it builds up to more simulation exercises.

These exercises do not involve the scenarios that we want to study. They are other kind of accident scenarios or smaller parts of other types of scenarios. So they don't get any specific training on the scenarios but to familiarize them with the interface and the process.

Very briefly about the design for the study, I said the purpose of this study was another and this empirical investigation, but the study contained two types of scenarios, accident scenarios, SGTRS and loss of feedwater. And the study was actually focusing on how scenarios can become complex for the crew to handle. And the design was such that we had what was called a base case for the scenario. In the SGTR it's a -- what we call a clean SGTR. There is no other complications added to the scenario than --

As part of the complex case there are some more failures like another event starting before -- I'll come back to that later.

CHAIR APOSTOLAKIS: And there is a procedure for both the base case and the complex.

MR. BRAARUD: There is procedures that apply to both cases.

CHAIR APOSTOLAKIS: Okay.

MR. BRAARUD: Also a detail, these scenarios, they are balanced on the first rundown, so there are different orders for different crews.

MEMBER STETKAR: Did the crews -- I hope not but did the crews know that they were running two ruptures and losses of feedwater?

MR. BRAARUD: No, they didn't know anything about that.

MR. BLEY: And if I understood -- I think some crews may have seen the complex case first and other crews may have been the base case first.

MR. BRAARUD: No, they didn't know the balances or scenarios, so you don't run the complex always after the base because then you will have some learning effect that would make useless all the way around.

MEMBER STETKAR: But did you allow any

crews just to sit there for six hours with nothing happening?

MR. BRAARUD: No, we didn't have time for that.

MEMBER STETKAR: I'll get to my point here. If the crews knew you were running simulator experiments to evaluate operator performance, and the crews are not stupid. They know that they're there for a purpose. And they are all familiar with complex operational events in nuclear power plants, those events tend to be tube ruptures, ATWS and complete loss of feedwater if it's not a station black-out.

Did you at all consider the fact that the operators had this knowledge and that they may have been predisposed to look for things like a tube rupture because that -- or an ATWS. Now, it's pretty obvious that I don't have an ATWS if the reactor trips when it's supposed to trip. So if I don't have an ATWS and I don't have a loss of all feedwater and they're monitoring my performance, then maybe I have a tube rupture. So maybe I should be looking for a tube rupture. Did you try to correct for that? That's the reason for asking did you let any operating crew just sit there for six hours and do nothing?

MR. BRAARUD: That is a good point and in some other studies we have actually put in scenarios where nothing happens, reactor normal operation to take away --

MEMBER STETKAR: But you didn't try to do that here.

MR. BRAARUD: Not in this study. It was not possible in the schedule, but there is one important operation and that is in the complex SGTR, as several of the crews have problems and there are no indications that they are actually suspecting an SGTR.

MEMBER STETKAR: But there's some indications that they needed an excuse to go to E3 and they found it.

MR. BRAARUD: Yeah, they found it after awhile but several crews has used quite a long time and there was no indication that they had some idea.

CHAIR APOSTOLAKIS: Are crews trained on all of these scenarios routinely?

MR. BRAARUD: Yeah, in their --

CHAIR APOSTOLAKIS: Including the complex scenarios?

MR. BRAARUD: Not the complex.

CHAIR APOSTOLAKIS: Not the complex, never?

MR. BRAARUD: Not this specific

configuration or scenario.

MEMBER ABDEL-KHALIK: In some plants they are. My question about the complex scenario, for example, how is that selected? Why not, for example, make the scenario in such a way that the indications would be counterintuitive. Like if you have a loss of one reactor coolant pump in a three-loop plant, and a steam generator tube rupture in a different loop, that would be an event that I'm not sure any of these crews would recognize. And they would have to follow the procedure by the -- you know, step by step to get to where they're supposed to go.

MR. BRAARUD: That's a good proposition for a complex scenario, but in this one we choose one that has to do with the indications of the SGTR. That was some of the research questions for the original study, how a situation becomes complex when you have previous information that point to another even than the actual event coming.

So, other studies could look at other complexity issues, like you suggest.

MR. BLEY: Per, can I say something? The thing -- the discussions that led to this began many years ago and when many of us doing HRA were trying to talk with folks at Halden, "About how could you do things at the simulator that would help us". And one of the things we had talked about was we're really interested in difficult scenarios, maybe with some kind of masking.

So they came up with this one, trying to find something that would make this more challenging and test some of the things we try to do. So there are maybe lots of those but this was one they could live with.

MEMBER STETKAR: This is a pretty good scenario, by the way.

MR. BLEY: Yeah, and you could almost have a logic for --

MEMBER STETKAR: Because of both thermal hydraulic response for cool-down and the masking of, you know, elimination of the rad monitors.

MEMBER ABDEL-KHALIK: Well, you will always know which generator has failed even in this scenario. Eventually, right, but what I'm talking about is a scenario in which they have no idea which generator has failed and they will actually have to follow the procedures. That eliminates the variability of having people jump from one procedure to the other.

MR. FORESTER: And that may be a very good suggestion for our next scenario.

MEMBER STETKAR: With radiation monitors, 16 monitors, you'd know. You have to eliminate the N-16 to eliminate the generate -- a lot of the generators.

MR. FORESTER: Keep in mind, we're not testing the crews. We're really testing the HRA teams. They're just trying to predict what the crews are doing or what they're going to do.

MEMBER ABDEL-KHALIK: Right, one of the variability is jumping in the procedures. And you want to -- if you want to eliminate that, you want to make sure that everybody follows the procedures.

MR. FORESTER: We want the teams to predict what the crews are going to do, given whatever the scenario is but there may be better scenarios. I'm not disagreeing with you but our emphasis, again was on what the HRA teams would do given the scenarios and the context that the crews were operating at.

MEMBER STETKAR: My only question was, you know, is -- because there are a limited number of scenarios that you're running obviously, because you have a limited amount of time, and the operators knew that the purpose of this whole exercise was to evaluate somehow their performance, I don't know how much discussion there was with the particular crews, was there anything about the way the experiment was designed in -- for this particular exercise that -- the tube rupture, was there any way that the experiment was designed that might predispose the operators to be looking for a tube rupture more carefully than they would if they just had a blank slate. You know, after the plant has been operating for 12 years without a reactor trip and nobody has ever seen anything except simulator training, that's --

MR. BRAARUD: I don't think there was anything that could lead them to suspect that we had SGTR scenarios.

Okay, brief illustration of the scenario and there was two versions of the SGTR. Quickly the base scenario, they have picked up a tube rupture, all activity indications worked as normal and they would get alarms and indications as expected.

While in the complex scenario there was initial a steam line break that actually isolated two of the expected activity indications, these red crossing. In addition, there was a failure with one of the measurement (phonetic) from the sampling from the ruptured SG.

MEMBER STETKAR: I'm sorry, Per, I need to ask one more question about this plant. It sounds like my reading of the scenarios, this plant does not have automatic auxiliary or emergency feedwater level control; is that correct?

MR. BRAARUD: I'm not too sure.

MR. BLEY: It sure looks that way from the plots.

MEMBER STETKAR: Because if it did, you wouldn't necessarily see a rapid change in the steam generator. You'd see a difference in loop-to-loop feedwater flows but not in level. If it's a fixed feedwater flow plant, then I can understand rapid changes in -- I'm just trying to understand the scenario and factors that might effect people's interpretation of the scenario.

So you do know, does the simulator have auxiliary feedwater level control or is it a fixed flow?

MR. BRAARUD: I think it is a fixed flow and that was -- the crew typically manually works with the control later in the scenario.

MEMBER STETKAR: Yeah, okay, okay.

CHAIR APOSTOLAKIS: So would you explain this figure?

MR. BRAARUD: Yeah, the figure, this is for the complex scenario. They first have a steamline break.

CHAIR APOSTOLAKIS: Oh, you have a steam line break.

MR. BRAARUD: Steam line break, yeah, first in the --

CHAIR APOSTOLAKIS: Where?

MR. BLEY: Outside of the MSIV.

CHAIR APOSTOLAKIS: Okay.

MR. BRAARUD: Off of the main steam line valve.

CHAIR APOSTOLAKIS: Okay.

MR. BRAARUD: Which actually lead to an automatic isolation of the steam line and that takes away two of the activity measures; one in the steam line and one from turbine. And there's one additional failure, from the sampling of the ruptured steam generators, failure with the measurement. That leads to no activity also from this one.

MR. BLEY: So there's no bad monitoring.

MR. BRAARUD: There's actually one failure and one event, previous event, that leads to this situation.

MEMBER STETKAR: Okay, why didn't you

break the steam line inside the containment? Same effect on secondary -- much more complicated in terms of what the operators are seeing. They would have never found the tube rupture.

MR. BRAARUD: That is also a very good proposal. I think we have two very good proposals for scenarios when we investigate --yeah, so this is the basics based on the complex.

MEMBER STETKAR: Too busy doing a lot of other things.

MR. BRAARUD: They're doing a lot of other --

MEMBER STETKAR: It's a real mess.

MR. BRAARUD: Okay, very briefly, this can also be operated by a typically event tree.

MEMBER STETKAR: I'm going to have to stop you here.

CHAIR APOSTOLAKIS: Let him explain.

MEMBER STETKAR: I'm sorry, Mr. Chairman, but I need to make a couple of points here. This event tree in the report was characterized as a typical event tree for a tube rupture. In my experience, this is not a typical event tree for a tube rupture. The reason for that and it's very pertinent for this particular exercise and it's an extremely pertinent concept for HRA in general.

You have defined the key operator response in this model as identifying and isolating the ruptured steam generator. If the operators do not identify and isolate the ruptured steam generator, you basically say they will never initiate a rapid cool-down, they will never initiate a primary pressure reduction.

They can survive if for some reason they make up to the RWST but the identify and isolate is required for success. I will not argue with you that identifying the fact that I have a steam generator tube rupture event is an important cognitive response. However, the requirement to both identify the fact that I have a steam generator tube rupture event and identifying the fact that this particular steam generator is ruptured and I must isolate it, are two completely different types of operator cognitive responses.

By defining a particular human failure event that requires the operators to not only identify the fact that they have a steam generator tube rupture event, and to specifically identify the correct steam generator and to isolate it, is a modeling decision that effects the definition of that human failure event within the scope of the PRA and which will have a very, very important effect on the HRA.

As you've seen, many, many people spent a lot of time evaluating the human error probability for failure to perform the implementation of isolating the steam generator which, if that's the way you define the human failure event, that's fine, but in fact, in that little event tree that you had as an example, in the real world, identifying and isolating the ruptured steam generator is not the key event that you've made it in your model.

In fact, identifying the fact that I have a steam generator tube rupture is important. Whether or not I isolate the specific steam generator can effect further event progression, but it does not preclude my cooling down and depressurizing. It does not preclude other event scenarios that are, indeed, possible and indeed, perhaps quite likely in your model.

So the point is how did you make the decision to define this particular human failure event because it is -- by including the identify a particular steam generator and isolate that steam generator in your success criteria you have now made many modeling decisions and assumptions that will have an extremely strong effect on perhaps some human reliability analysis.

CHAIR APOSTOLAKIS: So can you explain that in terms of the event tree? The first human failure event there is steam generator isolation. Do we -- have they assumed that they know which steam generator to isolate and you are saying that doing that is not the --

MEMBER STETKAR: This event tree, by the way, is different from the one that's in the report. I just noticed that. This event tree actually solves some of the questions that I raised regarding the report but this event tree is different from the one in the report. So this event tree has a different philosophy or a different interpretation of the event than what's in the report.

CHAIR APOSTOLAKIS: But if we look at the first event they have there, steam generator isolation, they have assumed in this even tree that they know which steam generator to isolate.

MR. BLEY: No, the --

MR. BRAARUD: No, the model includes diagnosis before, also this steam generator isolation part includes the diagnosis part or that they have a tube rupture and it's based on the procedures, diagnosis and entering a procedure for identifying which steam generator and isolating.

CHAIR APOSTOLAKIS: So this event includes the diagnosis of which --

MEMBER STETKAR: It includes, importantly, it includes the manipulation of isolating the steam generator. This says that -- and this event tree is logically similar if I look at it. If I correctly identify the fact that I have a tube rupture event, a tube rupture event and I correctly identify the ruptured steam generator, however, if I fail to close a particular valve, then in this event tree, I am not given the opportunity to cool down the secondary side, to depressurize the secondary -- the primary side or to save the plant through the normal method. This model has a particular modeling assumption built into it and in experience, looking at HRA, what we find is that these types of assumptions about modeling and defining the human failure events, in many cases have a much broader variability than a particular HRA methodology for evaluating the human failure event once it's defined.

In other words, the biggest problem in HRA is not necessarily the calculator that I use to calculate human error probability. It's how do I define the structure of the logic model and define those events?

CHAIR APOSTOLAKIS: So --

MEMBER STETKAR: And this is a prime example of that, that's why I wanted to bring that up.

MR. PARRY: Is a sense, you're really arguing about what goes on, on the failure path of that particular HFE.

MEMBER STETKAR: Well, absolutely but that defines -- your Human Failure Event success criteria require three things. I must identify the fact that I have a steam generator tube rupture event which sets me on a particular trajectory.

MR. PARRY: Right, right.

MEMBER STETKAR: I must identify the particular steam generator that is ruptured and I must successfully complete the isolation tasks within some time period.

MR. PARRY: Right, right.

MEMBER STETKAR: Those three things are required to be on the success path.

MR. PARRY: Right, and -- but that's -- if we're only focusing on that particular HFE, then I don't see any problem with that.

MEMBER STETKAR: Not for the purpose of

this study. My comment --

MR. PARRY: For this study, right.

MEMBER STETKAR: My comment is that this -- I wish you hadn't put the event tree in there, but as long as you put the event tree --

MR. PARRY: It's a real event tree.

MEMBER STETKAR: As long as you put the event tree in there, it brings up a very, very important point and in fact, the definition of this Human Failure Event apparently created some problems for some of the HRA teams because they were arguing about the timing which is something else that I'll bring up.

MR. PARRY: Which is a different issue.

MEMBER STETKAR: That's a different issue.

MR. PARRY: Right, but given certainly at least for this pilot study which is looked at the first event in that event tree, I think your point is not -- doesn't really pertain to that other than the aspects of the timing, which --

MEMBER STETKAR: It may pertain for the larger purpose, though, that George was talking about is where are we focusing effort in terms of HRA? Is it -- should it be focused on comparing different calculators for HEPs or spending as much or more resources on defining and modeling the Human Failure Events?

MR. PARRY: Yeah, that's a different issue, I think, I mean, particularly in the context of this study, but Vinh tells me that that's a real event tree from real PRA study.

MR. BLEY: The one that's in the report?

MR. PARRY: The one that's in the report.

DR. DANG: And we did not make that event tree in the study. So we're using -- I think one point that you make that's very valid, sorry, is that maybe we should be careful not to say that it is typical. I do not know whether it is typical.

CHAIR APOSTOLAKIS: But, John, how would you change the event tree?

MEMBER STETKAR: I don't want to get into the modeling, the details of the modeling. I had to get into it a little bit make my point about the modeler's assumptions. It's always said, "Well, the event modelers or the so-called systems analysts define the Human Failure Events", and then the poor HRA people analyze the Human Failure Events within the context. And everyone focuses on the second half of that process about the differences in the calculators.

MR. PARRY: But that was one of the

intents of looking at this was the calculators, because you realize that --

MEMBER STETKAR: That's true, that's one part of -- absolutely, absolutely.

CHAIR APOSTOLAKIS: Again, let's -- for my own benefit, steam generator isolation includes the cognitive part?

MEMBER STETKAR: Yes.

CHAIR APOSTOLAKIS: It does? What John says, you know, not a steam generator, but this steam generator.

MEMBER STETKAR: I mean, there's two levels of cognition in a tube rupture event. The first cognition is do I have a tube rupture or not?

CHAIR APOSTOLAKIS: Right, then which one.

MEMBER STETKAR: The second part is which one.

CHAIR APOSTOLAKIS: And that's all in there.

MR. PARRY: It's all included in the --

MEMBER STETKAR: And the manipulation, closing it all down.

CHAIR APOSTOLAKIS: And the manipulation, yeah, all three, everything. Okay.

DR. LOIS: I believe we have to clarify.

The scope of this study is pretty much quantification. The capability of a method to understand the scenario and identify the various parts that the operators could take was not part of the study. We were focusing more into assuming that now you have this human action. This is defined, this is information associated. How would you analyze that specific human action as opposed to use your method to understand the scenario, determine your human actions and then analyze them. It's a different study.

MEMBER STETKAR: I understand that, but the point that I -- we'll get to something that's a little bit more relevant in a couple of minutes here, I hope.

CHAIR APOSTOLAKIS: But let me ask you one other question. You said that this event tree sends a message that if they don't isolate and if they don't cool down the secondary site, they will never depressurize. Is that what you said, whereas they could --

MEMBER STETKAR: This -- HRA in an analysis space, if the operators fail to close, for example, one local manual isolation valve in the auxiliary building, they will fail and they are never given the opportunity in this model to cool down the secondary site or depressurize. Now, that is a very, very, very important modeling assumption.

CHAIR APOSTOLAKIS: And that refers to the structure of event tree, right?

MEMBER STETKAR: That's absolutely true.

CHAIR APOSTOLAKIS: Okay.

MEMBER STETKAR: That's a logical concl --

MR. PARRY: It's not the HFE that we're interested in.

CHAIR APOSTOLAKIS: Right.

MR. FORESTER: We only had them quantify events on the success --

MEMBER STETKAR: That's right, I'll get to the -- I'm going to get to the second part of the concerns.

CHAIR APOSTOLAKIS: Shall we move on from this event tree then?

MR. BRAARUD: This just gives an overview of the HFEs you're talking about, for the base and for the complex scenario.

CHAIR APOSTOLAKIS: Why do you call these major?

MR. BRAARUD: It's -- yeah, it also actually describes a major part from the procedures, so it's actually the HFEs and it's the major unit performance. It's a little misleading title. But the point is, that's for this first pilot phase. It's only the first HFE and that is focused and that includes the diagnosis and identification and the isolation in this case.

MR. FORESTER: That's the word analyze but the rest of the data was collected on those other events.

MR. BLEY: The HRA teams analyzed all of these.

MR. FORESTER: Right.

MR. BLEY: The comparison has only been done on the first one.

MR. BRAARUD: There is also some more to the end of one of the scenarios. This is actually HFE as such, that's maybe why the titles says "Major actions".

So this actually, this HFE and the prime success material that is based on the diagnosed that they have SGTR and therefore, enter the procedure of SGTR reaches the E-3. And that they isolate the ruptured steam generator and also about they stop the feed to the ruptured steam generator so they don't over-fill the steam generator. This was based on the procedures and also based on the home plant's task analysis of important tasks.

And there was developed a timing criteria for this HFE. That was based on what one would expect from training, that the crew has a focus and that they have the possibility for letting out active steam, so that they have a pressure to work quite swiftly, quite quick, and it was based on the procedure steps or the major actions involved in reaching this criteria.

CHAIR APOSTOLAKIS: This 20 minutes in the discussion in the report puzzled me a little bit.

MEMBER STETKAR: Let me follow up on this, George.

CHAIR APOSTOLAKIS: Okay.

MEMBER STETKAR: Why didn't you use a functional plant response timing success criterion as is the recommended practice in PRA? In other words, the time available for the operators to perform an action is determined by a change in plant conditions that essentially requires either -- either will initiate some automatic response or requires another operator evaluation of the scenario. Twenty minutes is an arbitrary time here. It's an estimate of the amount of time that you, as an outsider, expect the operators to spend going through those procedure steps successfully.

Well, if I'm driving my car and a small child darts out into the street 100 meters down the street and if I go through all of my training and procedures and say, well, I expect that I will stop the car within two meters of the child, that's a certain time. I suspect that's not the way the world really works.

So why didn't you actually use a functional plant response time window, which for this particular action, and this is where I'm getting to defining the action, in the analyses that I've seen that I'm familiar with, it really doesn't make any difference when the operators isolate the steam generator. And in fact, in many plants, the operators are trained that that is a much secondary function compared to the other more important functions of cooling down and reducing pressure. So by arbitrarily assigning a 20-minute time window that's not based on plant thermal hydraulic response to an action that may be assigned very low priority by many operating crews, at least outside of Sweden, you are very strongly biasing perhaps the results from some human reliability analyses that look at relative timing, time reliability correlations, for example, or effects of operator priorities as reflected in the emergency operating guidelines in the background documents and those guidelines, in fact, de-emphasize isolation of the ruptured steam generator in favor of doing some other things like cooling down and depressurizing.

So I'm really curious about this time window because it's -- it seems very arbitrary and extremely artificial and in fact, contrary to recommendations.

MR. PARRY: I think in some sense -- may I make a response? I think in some sense you are right. From a PRA perspective, you wouldn't define the time window this way. And I think we've taken that as one of the lessons learned for future simulator exercises that they ought to be -- the timing issue particularly, ought to be thought differently and I think it did effect some of the team.

MR. FORESTER: But on the other hand it was tied to isolating the steam -- we didn't want to overfill the steam generator.

MR. PARRY: But John's right.

MEMBER STETKAR: But cooling down and depressurizing stops you from doing that regardless of whether it's isolated. That's what the operators are trained to do. But in some sense success criteria HFE constrained the problem for the HRA teams. The more general problem event model, the second action that you have defined initiate the cool-down.

MR. PARRY: I can comment on the training valid in the sense numerics of HEPs, but more on the qualitative analysis that the human reliability analysts did, because the HEPs in particular could be, because of this somewhat artificiality of the success criteria and would not be relevant in the PRA context.

MEMBER STETKAR: As long as you're -- you're also sensitive even in a qualitative sense, if some of the HRA methods included something like a reluctance factor, if you can think of a reluctance factor, that the operators would be reluctant to perform this action because it might delay their more important action, which is initiate the cool-down. So there could be, you know, that -- even in a qualitative sense that understanding what some of the methods include, because of the definition and the artificial time for this action, and because time is a very important consideration, I guess we can all agree on a tube rupture event, that may have an influence even in the qualitative sense. When you compare the different events, you say, "Oh, these people made this decision in their methodology because they do include, you know, something, I'll call it reluctance factor for this particular action and that might be something that the other methodologies did not consider. But that's --

CHAIR APOSTOLAKIS: But there's another thing that is stated in the report in this context of -- I need some clarification on. It said, "We usually give the -- assign this time of 20 minutes or 30 minutes in other cases having in mind core damage", but the operator is not thinking in terms of core damage. It's thinking in terms of an intermediate event like filling the steam generator and so on. I'd like to understand that a little better and how it effects the performance.

I mean, why are we doing that in a HRA, I suppose and say, you know, 30 minutes and then you have core damage when, in fact, these guys are not trying to prevent core damage at that point. How does that effect the whole calculation, the whole performance, because, you know, if you start putting -- like in most HRA models, there is a stress factor, right?

And you assign the level of stress based on core damage when, in fact, the guy is not thinking in terms of core damage, then you may be completely off.

MR. FORESTER: But they should be able to predict that. If they're only thinking about they want to avoid -- they don't want to release, so the HRE teams have to understand that what the operators are trying to do is avoid a release and that's why they're trained to do this, isolate this thing within a certain amount of time. So our decisions there -- there are some problems there but it does reflect what their training is and the fact that they do want to avoid the over-fill because they don't want to release and that's part of their training.

CHAIR APOSTOLAKIS: But --

MR. FORESTER: But you're right, I mean, ideally, we maybe would look at avoiding core damage.

CHAIR APOSTOLAKIS: Well, we, as a PRA analyst, think that way but if you tell me that the operator is not thinking that way --

MEMBER STETKAR: Well, the key is the time windows are usually set based, as I said on --

CHAIR APOSTOLAKIS: Thermal hydraulics.

MEMBER STETKAR: Plant -- getting to, you know, what is the time at which the plant -- if the operators don't do anything, that the plant parameters will change either to a point where some automatic action occurs, which changes the nature of the scenario, or a different requirement for operator performance is involved, and that's not necessarily core damage. It's usually some intermediate condition.

CHAIR APOSTOLAKIS: Yeah, but --

MEMBER STETKAR: A pressurizer goes empty or a steam generator fills with water, or something like that. It isn't necessarily core damage.

CHAIR APOSTOLAKIS: But is it true though that when you do the HRA, and you consider those time intervals and you say there will be this kind of stress and this kind of this and that, people are thinking in terms of core damage but when the report says that the operators are not necessarily thinking core damage.

MEMBER STETKAR: No, if you do the HRA, you think within the context of the scenario that's defined, the amount of stress that would apply during whatever your perception of that time window will be.

MR. FORESTER: So this wouldn't be high stress like in a large LOCA or something like that.

MEMBER STETKAR: That's right.

MR. FORESTER: This is a different kind of scenario, but the HRA teams understand what the scenario is and they should be able to --

MEMBER STETKAR: Sure, the HRA --

MR. FORESTER: -- understand what the stress load should be given that context.

MEMBER STETKAR: Should define stress within the context --

MR. JULIUS: This is Jeff Julius, the way the event tree is structured. So depending on how the event tree is structured, you might have a timing factor that, you know, for that particular progression, introduces this thing at 20 minutes and again, that's not seen or felt by -- that doesn't mean it's an influence on operator response. It just means reflects the way the model is set up.

MEMBER STETKAR: I'll give you an example that for example, in this particular -- let's take the base case tube rupture, the level of stress for an operator may not be very high in the first 40 minutes of the scenario, let's say, because the plant response is relatively slow. Now, if the operators over-fill the ruptured steam generator, and they have a release going on, and now they suddenly recognize that they've over-filled the rupture, at that point, the stress may increase dramatically or to some extent. The time window to core damage really hasn't changed all that much, but the operator stress, how they react to that situation, may change at that transition point.

CHAIR APOSTOLAKIS: But this is what happens. I'm asking how it's modeled. I agree that this is what happens in real life.

MEMBER STETKAR: In an HRA --

CHAIR APOSTOLAKIS: Let me ask you this; why then if this is the case, you would use this concept of the intermediate event? I mean, if you --

MR. FORESTER: The what?

CHAIR APOSTOLAKIS: Intermediate event. That's what the words are in the report. If the operator indeed knows or the evaluation is done for this particular event, without really thinking in terms of core damage, there is no need to introduce this concept of intermediate event. That's why I'm asking for clarification because I believe you say there in Section 2.5.4, 2.5.4, I think, I don't have the report in front of me, that --

MR. FORESTER: I'm not sure where you're referring to. There's intermediate in the sense of maybe there are other events that are coming out but --

CHAIR APOSTOLAKIS: I think it says that the operator is not thinking in terms of core damage. It's thinking about, you know, maybe a minor release or so. But then you make a contrast between that and what the PRA analyst is thinking, which is core damage. I don't know why you make the distinction. That's what confuses me, because if indeed, as you said, the operators are thinking in terms of this particular event and they don't really think in terms of what will happen later, then there is no need to talk about this.

MR. FORESTER: No need to talk about --

CHAIR APOSTOLAKIS: Intermediate events, that's why it caught my attention. The message you were sending was that the PRA analyst thinks in terms of A, what is the operator, in fact, is thinking in terms of B. And I thought that --

MR. FORESTER: Every analyst ought to be thinking about what it is the operators are thinking about because that's what they're trying to --

CHAIR APOSTOLAKIS: Right, and you make a distinction between the two thinkings, that the operator is thinking one way. Maybe I can try to find the exact --

MEMBER STETKAR: I think I've found the paragraph, George.

CHAIR APOSTOLAKIS: Okay.

MEMBER STETKAR: You can look at it over

lunch. I have it.

CHAIR APOSTOLAKIS: Yeah, we can do that.

MR. BRAARUD: But I think one important point though is that this time criteria is maybe somewhat arbitrary related to the plant response. Of course, time is related to the importance of letting out the steam but the important thing is that this was defined as an exercise for the HRA team to analyze human performance.

So it's still -- I know that HRA analysts may have in the back of their heads, some tendencies of thinking about the scenario in terms of how they have defined it for other purposes but in the information package, it's well-described what is the purpose of the procedure, what is the -- what you think are the goals and based on their training and procedures, and for the purpose of comparing similar data with HRA analysis, I think this is useful.

MEMBER STETKAR: I would only caution you that when you do the comparisons then of the HRA method, I don't care so much about the results, the numerical results, but the -- even the qualitative evaluations, that you try to take a bit more of a holistic look at the relative evaluations across the whole sequence of events because, for example, if HRA

Team A assigned a relatively optimistic assessment for identify and isolate, and a relatively pessimistic assessment of start the cool-down compared to Team B which may have reversed those, pessimistic for identify and isolate but optimistic for cool-down, there may have been real qualitative reasons for doing that, that this reluctance factor, if you want to call it that, that Team B who is pessimistic for the isolation may have been focusing more on the task of initiate the cool-down.

Team A that was, perhaps, optimistic may have said, "Well, they spent so much time successfully isolating that they didn't have enough time left to start the cool-down because of the way the actions were defined. So that when you're comparing the different methodologies to understand what the HRA teams were thinking about, you can't necessarily just do it, especially in this case because of the way that this HFE is defined and the time that is used to define success from the simulator, you may have a biased interpretation about the relative benefits and drawbacks of some of the different methodologies if you just constrain your evaluation to this one particular action, because it is -- the action, as it's defined is a relatively artificial action and the time is certainly artificial.

MR. FORESTER: There were some teams that felt that way.

CHAIR APOSTOLAKIS: Well, any time you are at a particular note in the event tree, you are assuming that you do nothing, there is a certain time until core damage occurs, and that time of course, changes from node to node. Then there is another failure, that time changes. So what matters, I think what you're saying is, that you can't really specify a time for one particular event. What matters is, perhaps the sum of these times that it takes them to do it.

So if you do something very quickly now, you have more time to do something later.

MEMBER STETKAR: In a tube rupture, without, you know, it's noontime, George, so you have to be a little bit careful about time. In a tube rupture event, one of the things you're concerned about is over-filling the ruptured steam generator.

CHAIR APOSTOLAKIS: Right, right.

MEMBER STETKAR: That is a time at which the fundamental response of the plant changes. It's not core damage. It's just the change --

CHAIR APOSTOLAKIS: No, okay.

MEMBER STETKAR: Okay, there are a few things that need to be done to avoid that condition. Now, how those things are modeled in a particular PRA, fault tree or event tree or whatever, defines the Human Failure Events that the HRA analysts must evaluate. The real time window is to avoid over-filling the ruptured steam generator. Now, that is a time window.

The intermediate things, what they do in between, there are different activities, different buttons that they need to push, if you will, to accomplish that eventual goal.

CHAIR APOSTOLAKIS: So the sum of the --

MEMBER STETKAR: But in a sense, you could model it as a single action, do the operators prevent over-filling the ruptured steam generator but that's rather vague. So a lot of the detail that we're talking about is how is the logic model structured and how are those actions defined? As soon as you define intermediate actions, you have to provide a context for them which is both scenario progression and timing, which is part of the problem that we're struggling with here.

CHAIR APOSTOLAKIS: The sum of the times to achieve these actions should be less than --

MEMBER STETKAR: Absolutely.

CHAIR APOSTOLAKIS: Okay.

MR. PARRY: Can I try maybe and bring us back a little bit on track because I thought the purpose maybe of this report is to demonstrate that this methodology of taking simulator results and comparing it to HRA analyses could give us insights into the various HRA methods. You're right, John, maybe these scenarios were not defined as well as they could have been, but given their imperfections and we understand that they are imperfect, can we still get insights into the strengths and weaknesses of the various HRA methods, which I think is what we're trying to do?

MEMBER STETKAR: I agree. I think it's -- it's a problem of the devil is in the details often. I was surprised -- I wasn't surprised you selected tube rupture because that's a great -- you know, and that's what you had -- it's a great example and that's what you had the team's performance for. I was very surprised that you selected this particular Human Failure Event in lieu of, for example, operator starts the secondary cool-down.

MR. PARRY: Because I think that's the way they're trained. My understanding is that part of this was chosen -- and I wasn't involved with defining these scenarios, so I'm not sure, but I think that's the way the operators are trained.

MR. BRAARUD: Yeah, that's the way the procedures are laid out.

MEMBER STETKAR: That's the way the procedures are laid out but --

DR. LOIS: And trained.

MEMBER STETKAR: Well, but does that mean that if I do not close this one valve, the operators will not depressurize the secondary -- will not cool down the secondary side?

MR. PARRY: No, but --

MEMBER STETKAR: That's what --

MR. PARRY: We're really looking at the success pathways.

MEMBER STETKAR: Then why didn't you select the initiate secondary cool-down because it avoids all of the complications of did the operator close this valve and the complications that many of the HRA teams focused all of their effort on the implementation part of that particular HFE?

MR. PARRY: We still would have had to have tripped the reactor.

MEMBER STETKAR: That's okay.

CHAIR APOSTOLAKIS: Let me interrupt here for a moment. Are we still doing Roman III on the schedule?

MR. PARRY: Yes.

CHAIR APOSTOLAKIS: So we are two hours behind. We are losing one member Dr. Abdel-Khalik at 3:30. John, are you leaving early?

MEMBER STETKAR: I'm leaving tomorrow, I hope.

CHAIR APOSTOLAKIS: You're leaving tomorrow. Dennis? Oh, Dennis is not -- we are losing also, I think the visitors at some point, at 4:00 o'clock?

DR. LOIS: Yes.

CHAIR APOSTOLAKIS: The visitors are not making any presentation in the afternoon.

DR. LOIS: Yeah, the assumption was it's -- if you look at the afternoon, V is again, two people, Per Olivind and Salvatore.

CHAIR APOSTOLAKIS: Yeah, but that will be early afternoon.

DR. LOIS: Okay.

CHAIR APOSTOLAKIS: Yeah, so the rest of us can stay a little longer? Yeah, because we can't stop this. It's important stuff.

MR. BLEY: I think for input on the schedule, I probably made most of the points that from my side that would require a lot of discussion so I can probably just sit here and listen.

CHAIR APOSTOLAKIS: John, are you leaving early? Oh, you're going tomorrow.

MR. FORESTER: Not me, I'm going tomorrow, yeah.

CHAIR APOSTOLAKIS: Okay, all right, so then we'll go out of our way to have Per and Salvatore -- they go before you or is there a logic to this?

MR. FORESTER: I can go after them.

CHAIR APOSTOLAKIS: You can to after them? Okay, so we'll go with you next which will be after lunch and then we'll stay here and have a nice evening. Per, can you finish this quickly?

MR. BRAARUD: Yeah.

CHAIR APOSTOLAKIS: Okay, so let's move on then to 17.

DR. PRITCHARD: Preferably, the HFA for us is the complex scenario is defined in the same way and they have the same metrics that we discussed so we don't have a long time based on the same principal training and so on, same things we discussed. I can say very briefly about the whole -- how well the procedures matched the scenarios which I would go over.

For the base scenario there is a good match between the procedures and the plant allotment. They have radiation activity indications as expected and use that as a key step to transfer -- to diagnose and to transfer to the procedure. While in the complex scenario, if we obtain full activity indications, the crew will not find this information in the step they usually diagnose as the job. So they will then continue in the E-0 which is the diagnosis procedure and depending on how they run the scenario, they will either have conditions for transferring to a procedure or stopping the safety injection or continuing further in the E-0.

There will -- in both these parts there will be the support for diagnosing that they have a SGTR. But it's more uncommon, more unfamiliar for the crews in the complex version. Since that's the main message. So I don't think I will got through the details. That will also come later when --

CHAIR APOSTOLAKIS: Okay.

MR. BRAARUD: Very briefly, I can say that the simulation went quite well. It's almost a question when you have a simulator. It can be a technical problem, problems with the storage of data and such thing and generally, this worked out well with some minor problems. Also important that the crews be interviewed and they reported that they thought the scenarios were all realistically, comparable to how they would operate such a scenario in the training situations, this posed a little bit that the interface is not a very important factor in this. The crews didn't experience this as different from the training so much.

Also the base and complex scenario was handled quite differently by the crews so that it what I mean that the planned effect you could actually see quite a lot of interesting things in the --

MEMBER STETKAR: The crews at their home plant have interactive display, digital displays of their plant? I mean, they can call up parameters on digital displays?

MR. BRAARUD: Yeah, they have some systems are on the computers at the home plant. They're not completely on --

MEMBER STETKAR: Not completely, so it's not quite as integrated as the simulator. Okay, go on.

MR. BRAARUD: This is the end of this

presentation.

CHAIR APOSTOLAKIS: Very good. So we're ahead of schedule. Okay, we'll be back at 12:55, one hour.

(Whereupon at 11:54 a.m. a luncheon recess was taken.)

CHAIR APOSTOLAKIS: So the next presentation is by Mr. Salvatore Massaiu.

MR. MASSAIU: (Pronouncing) Massaiu.

CHAIR APOSTOLAKIS: (Pronouncing) Massaiu. And Parry again. Where's Parry? Oh, you are not presenting?

MR. PARRY: Afterwards.

MS. LOIS: And in that session also there's a name missing, which is Vinh.

CHAIR APOSTOLAKIS: Not again.

We are abusing you, Vinh. We shouldn't do that.

Please go ahead.

HAMMLAB DATA ANALYSIS AND RESULTS

MR. MASSAIU: I will present about the data we collected in HAMMLAB from the simulator.

CHAIR APOSTOLAKIS: We can't hear you very well. Can you move the microphone closer to you?

MR. MASSAIU: This one?

CHAIR APOSTOLAKIS: Whichever one.

MR. MASSAIU: So the data we have, we analyzed them for the study. And part of the results.

The results which I will present are the performance in terms of response times, times to isolate, and so-called operational summaries, which is kind of aggregated stories of what we have seen the crews in the two scenarios.

And then someone else will present about the PSF, derivation of coefficient. And some details of what we've seen as well.

So performance times, for isolation, the two scenarios. Especially saying something about variations of scenario – of procedures, regulations in the complex case, so much variation at least in the base case. And the PSFs we take in in our presentation, Vinh's presentation.

So that's our sources. When we run simulations, we have audio-video recordings of everything. There is an expert, process expert, who his also commenting on what the operators are doing.

When we are finished we interview the crews. And we went through each phase of these scenarios specifically.

Everything is logged into the simulator

log system so we have all process parameters and all interactions where the interface is logged.

We have a so-called all-process system, operators perform an analysis system, which is a task analytic instruments. We have a list of actions the crew is supposed to do, and then you check when to do it, just to help the analysis.

We have also a rating of the performance. This is the expected observed, and was also given a judgment of how well the performer -- for example, how well you perform isolation, just as a quick reference for later analysis.

And then when the simulation is finished, and then which scenario the operators are also asked some questions. This is the observer – the operator PSF rating.

And this experiment we also had four operators, normally senior shift supervisor. So it's not -- it was just sitting there in the control room beside the other, not interacting with the other, and just making some evaluations on his own.

And then we asked some questions, background questions. How many years have you been working and so on and so forth to all operators of the experiment.

CHAIR APOSTOLAKIS: It is not what – the same thing –

MR. PARRY: It is not.

CHAIR APOSTOLAKIS: Is this screen working?

Okay, I'm sorry.

MR. MASSAIU: So all in all, 14 views, two scenario types, steam generators and loss of feedwater, two variants per scenario, and normally about 1-1/2 hour scenario runs.

CHAIR APOSTOLAKIS: Mr. Massaiu, can you tell us a little bit about your background?

MR. MASSAIU: I have a master of art in philosophy and master of science in marketing research, one from Italy, one from Norway. So I'm Italian, I work in Norway.

CHAIR APOSTOLAKIS: So you are not an engineering psychologist? You are more on the statistical side?

MR. MASSAIU: Partly on the statistical side from marketing research, and partly on social work, philosophy.

CHAIR APOSTOLAKIS: Good, thank you.

MR. MASSAIU: In terms of the overall data analysis process we had the so-called raw data which I was describing on the video, so on and so forth, and which relates to these 14 crew level performances. And for each you can imagine that they have an operational story, how they operate it, and they have a special configuration of PSX which depends of course on the interaction of the system.

And you could also describe the difficulties they had in performing the scenario.

From this level we needed in this study to go to a higher level evaluation, because normally the HRA is not working on a crew specific level. So we had to put together all these 14 crews, the base case and the complex case, into two operational aspirations. So it's kind of, how did they operate overall, these crews, for each scenario? And what were the driving factors, as a set, for each of the two conditions.

MEMBER ABDEL-KHALIK: Just a question. The sets of three people remain together throughout this process?

MR. BLEY: The crew stayed, those three people did all the exercises together?

MR. MASSAIU: Yes.

MEMBER ABDEL-KHALIK: You didn't mix them?

MR. MASSAIU: No, no, the same persons were operating from the beginning to end of the scenario.

MEMBER ABDEL-KHALIK: Okay, so was there any attempt to sort of check for correlation in performance in the four scenarios?

MR. MASSAIU: Not in each stage, we didn't analyze that.

MEMBER ABDEL-KHALIK: In the sense that you have a particularly strong crew, and a – well, not quite as strong crew, and was the performance in the four scenarios correlated?

MR. MASSAIU: No, we haven't done an overall evaluation of that.

MR. BLEY: Are we going to?

MR. BRAARUD: We are going to do that. That is one interesting question. We are going to analyze that in another project, not this very investigation, but for the purpose of understanding how crew differences actually relate to how they operate.

CHAIR APOSTOLAKIS: When you say, another project?

MR. BRAARUD: That is the original project that this data come from.

CHAIR APOSTOLAKIS: So this is still NRC? The NRC is participating in that?

MR. BRAARUD: It's the whole department.

CHAIR APOSTOLAKIS: So we benefit from this?

MS. LOIS: Yes.

MEMBER ABDEL-KHALIK: But if there is a correlation in performance wouldn't that indicate a bias in the predictions of whatever models?

MR. MASSAIU: It depends what are you going to predict. But I don't think at the level which you have on the right in this table does matter.

MEMBER STETKAR: It would however say reinforce this notion that they do bring up in the report about the relative importance of crew-to-crew variability. If you saw a consistent correlation across the four scenarios, that would tend to reinforce the crew-to-crew variability as a strong input to HRA uncertainty for example compared to kind of random performance over the different scenarios.

So it could be an important input from that perspective, because that seems to be one of the messages from this phase of the study that there is measurable crew-to-crew variability.

MR. MASSAIU: Yes, it would be a stronger basis to evaluate the importance of that factor in the context in the set of the driving factors as I said, but on the other level I don't think it should be that important.

MR. BLEY: Salvatore, before you change, I wondered if I could try to answer your earlier question, and then other people here might add more. Maybe there is more coming in a later presentation.

I think this chart on the left hand side shows part of the reason why doing this at Halden is different than trying to do it at a simulator at a power plant. All of those instrumented recorded results are all linked together. So after the event you can go back, and you can watch exactly what they were doing, key it to what's in the logs, and really keep examining this scenario for a long time.

Plus the group of people they've assembled have a lot of expertise in how to set up the experiments to try to control for things like learning, and all those things they talked about earlier.

MEMBER ABDEL-KHALIK: But at the same time I recognize you know all the hardware and software needed to do this is not really that difficult to implement in a typical simulator in a U.S. plant.

MR. BLEY: There have been people who have tried that, and I think it's kind of expensive. It's not impossible, but how big an effort it would be, I don't know if anybody here can speak to that. I know when people were trying to do some of that before, it's not so easy from what I've heard.

MR. PARRY: EPRI did a lot of that back in the `80s.

MR. BLEY: But they still didn't have stuff linked back into the computer log and all that.

MR. PARRY: No, they had a video and some stuff like that, but not much else.

MS. LOIS: But Halden has offered to come to U.S. and replicate their studies in a U.S. setting.

The most important thing is their expertise to collect the data and analyze the data. It's a tremendous amount of resources. In this study we have at least four full-time personnel from Halden that is helping us out run this research.

MEMBER ABDEL-KHALIK: So that's the biggest advantage, the people who brought in the hardware.

MR. BLEY: I think it's both. Because although in principle it doesn't sound like a big deal, whose going to put together that finding? And you are not going to be able to tap into the plant computer, or probably even a simulator computer, and link up this kind of data so you can look at it six months later and really have it all tied together. I think that's a bigger deal that it sounds like.

MS. LOIS: Yes, it is. It is.

CHAIR APOSTOLAKIS: Okay.

MR. FORESTER: But it'd be good if plants had that capability.

MR. MASSAIU: So this was the goal. You have 14 crews, two scenarios, a lot of data, and then we have to go to a higher level when we summarize this information. Because normally it's the analysis level of HRA.

CHAIR APOSTOLAKIS: So what is the – what are we looking for? I understand we want to see what BSFs the analyst used, right? What else? I mean when you are doing the analysis, are you looking for something? Or are you doing various things and see what happens?

MR. MASSAIU: I can give you detail of what we're doing.

So the first step was to take all crews, see how they did perform, use simulator logs and other quantitative kind of data to see how – which were the fastest performing, and which were rated better than other crews.

And the second one was a selection of a subset of crews for in-depth analysis, and I will say how we did later I will come to explain what kind of in depth analysis we did.

We had to select crews, and we wanted to identify a mixture of both ends of the performance spectrum, mainly which we can call best and worst performance.

And we used two criteria for this selection process which was the time they used to isolate, and they level in their steam generator at isolation.

MR. BLEY: Can I ask another question? Because this is one I had at the workshop, and I wasn't able to find it in the revised report. It's two pieces.

Is somebody going to explain how you evaluated best and worst in terms of performance, and what the expert performance rating is. That's described in the report, it isn't described how it's done that I could find. It just says, this guy gets a five and this guy gets a three. It's hard to tell.

MR. MASSAIU: That's true, but because we didn't use that as a criterion for the selection. In the first report, yes, it was wrongly said that we used that criteria. But in fact the selection was done only by isolation time and level.

MR. BLEY: So it was strictly based on, a guy is good if he gets it done fast, and he's worst if he –

MR. MASSAIU: I'll show you a table where we have both time and level and the crews, and so we selected six crews in the complex scenario and three in the base case.

But we analyzed also other crews in addition which are not in the report, and which are not in the derivation of the PSS.

MEMBER ABDEL-KHALIK: So there is no measure for the potential of misdiagnosing the event? If someone is jumping around in the procedures for example, that is not captured anywhere?

MR. MASSAIU: It is captured in the finals, because – I will come to the next slide, how we analyze in depth each of the selected crews. And then you will have the entire story in detail of what they did during the part of the scenario analyzed.

So you get the kind of detailed information.

The point three that's in the slide, I think it's important, because we analyze a subset of crews. This was also a requirement at that time, the scale of this project, and the availability of resources.

But we analyzed when we wrote the report other crews in addition to see if we could confirm the premises which are reported as results from the analysis of the subset.

MEMBER STETKAR: Salvatore, this best and worst is also only in the context of this particular HFE?

MR. MASSAIU: Yes, it is .

MEMBER STETKAR: So for example if a crew took longer to isolate the steam generator, but performed – started the cool down more rapidly, they would get a poorer mark for this HFE?

MR. MASSAIU: Yes, that's right. At this stage we actually analyze all crews in the complex case, and I would say that we don't have – I can't remember any case of doing bad in the first one and do it very well later on.

MEMBER STETKAR: No, no, I was talking about actions during the sequence, regardless of whether it's the base case or the complex case.

MR. MASSAIU: Yes, doing very bad in the first sequence, and then improving on later events. I don't remember any striking case of that.

So of the selected crew we performed in

depth analysis, which is based on DVD review. So we took the selected rounds, we assigned three different reviewers, and afterwards, we checked the consistency of their reviews which were made.

The review activity is approximately one day, but of an hour recording time. And in the review we use performance ties, the use the operator performance scores. We use the observer log. We use all our lab data, but we actually understand the situation of course.

In this part of the study we only analyze isolation identification, and we didn't analyze at that stage at least the rest of the event.

We also used the interview that we had with the crew after the scenarios. And we completed the so-called H2 header, which is – which parallels the Form B, I think you had said something about, which is a classification of PSS according to a common format.

So also we from the experience side, we were trying to use the same classificational system which was difficult to exercise, but we did that. So in order to make the comparison of the experimental results with the predictions easier by using the same language.

But the main result of the DVD reviews is that the crew summaries, the so-called crew summaries. So what are the crew summaries? They are in three parts.

The first part is a short story of what happened. It is a timeline description in which we excerpts of communication, but we also had comments about good performance, trying to – not only to have a timeline of things, as has been said, but also try to explain what is the meaning of that communication.

Another section, we had a summary of the most influencing factors, which are classified according to four categories, which are, what we rated to be directly negative influence, negative influence being present, neutral influences meaning PSS from a list which we don't think were influencing anything but probably it's not a very appropriate name, and also, positive influences, that we thought some PSF did positively influence the crew.

CHAIR APOSTOLAKIS: So what's the difference between the first two?

MR. MASSAIU: The direct negative is something which influenced the performance of the HFE directly. So it showed a direct effect.

The other one is a PSF which we believe

was negative, but which did not influence the – the performance time. It did not have an effect on how fast they did isolate. They did communicate not that well, but still, the fact that they did not communicate that well didn't have an effect on how fast –

MR. BLEY: Is it that – I was reading it, and is it that it didn't have an effect – the way I read what it said was, the first one it was clearly observable that this caused them a problem, and the second one they couldn't really pin down how significant it was.

MR. MASSAIU: Of course, you need more than a performance actually if you have to do a strict evaluation of what is direct –

CHAIR APOSTOLAKIS: Why did you feel you had to identify it?

MR. MASSAIU: Yes, part of the reason is this, if you think that something is really observable which really had an effect on the timing of isolating, you say this is a direct negative.

If you see something which is negative – let's say that two crew members are not communicating very well; one is not responding to the inquiries, but yet the supervisor is deciding, okay, we have to transfer procedure, the time would be the same of isolating after the communication was improved enough.

MEMBER STETKAR: It could be something like they're not reading back something to each other but yet they are still accomplishing the goal?

MR. MASSAIU: You have two level of performance. One would be like a qualitative performance of the crew, how good they were. Another one is the quantitative one, how fast they were, that was the criteria we used to select the crews and also to quantify them to work with the HRMSs.

MEMBER STETKAR: I'm assuming that the crews were speaking Swedish, is that correct?

MR. MASSAIU: Yes.

MEMBER STETKAR: There have been other constraints placed on people at times.

MEMBER ABDEL-KHALIK: Was noise simulated in these scenarios?

MR. BLEY: Like a steam line rupture.

MEMBER ABDEL-KHALIK: Would that have an influence on the operators in your view?

MR. MASSAIU: I am not familiar with the setting of the plant control rooms.

MEMBER STETKAR: It might have in the second case, because there was some evidence that the crews, I think, contacted an external operator to verify that in fact the turbine building was full of steam or something like that.

So it might have, in terms of diagnosis –

MEMBER ABDEL-KHALIK: I mean that's a very loud event.

MEMBER STETKAR: Yes.

MR. MASSAIU: We did not simulate that in the simulator.

MR. BLEY: I don't know if we have any nuclear plant simulators that do that.

MEMBER STETKAR: Some – I've been in Europe, and some of them have some attempt at noise for steam relief. I don't know about steam line break. I've never run into one of those. But steam reliefs, yes.

CHAIR APOSTOLAKIS: Okay, let's go on.

MR. MASSAIU: So the final part of the summary is – the summary – of the observed difficulties had in performing the JFE, this part is also in the Form E. So again, to facilitate the comparison of results.

The summary of the most intensive factor is the data basis for the derivation of the PSFs which is the total of the next presentation -- the following presentation.

CHAIR APOSTOLAKIS: This evaluation was done by the three reviewers?

MR. MASSAIU: Yes.

MR. BLEY: Who are Hammlab people.

CHAIR APOSTOLAKIS: All part of the team? Okay.

MR. MASSAIU: Yes, so this is the table of all crews. We have on the left side the base case and on the right side the complex case.

So in green we have the so-called successes, meaning they isolated within the time frame which defined HFES on this side. And the red, we have those that did not isolate within that time frame.

And I have underlined in gray the crews which were selected for – in that analysis. But again we also in depth analyzed more crews, and so in order to be more confident on the conclusions we were making.

MEMBER STETKAR: To kind of follow up on what I asked earlier, I notice in the complex case, crew K, although they failed this task, for some reason had the lowest steam generator level when they isolated it – much lower than any of the others, which might mean that they were focusing on a different task.

VOICE: And crew B –

MEMBER STETKAR: And crew B that was successful, actually, overfilled the steam generator, as did some of the failures. I'm more focusing on the people who failed this task but might have been doing something better on another task.

MR. FORESTER: There was a unique aspect to crew K. I forget exactly what it was. They took an action that nobody else did.

MEMBER STETKAR: But then again, from what Said – if crew K – our – without divulging too much information are the crew labels in both of these tables the same? Is K the same crew won the base case and the complex case?

MR. MASSAIU: Yes, I think these two crews are those which normally trip the reactor in the complex case. And the crew B, the level and the isolation might also be a simulator problem, so that's also the reason why we didn't choose that one as the worst performing and we skipped to the next one.

CHAIR APOSTOLAKIS: This is a very interesting table. It reinforces another set of results of holding the view sometime ago, in a presentation by Per and Andre was. There is crew to crew variability. There is -- I mean, look at the times. And I'm not sure where modeling any of this, the HRA. You don't – you may want to call it crew to crew, or you may want to call the random variability of the time to achieve something.

Is that a major message? This is something important – especially look at the complex case.

MR. FORESTER: Yes, certainly some of the methods recommend looking at critical variability. I mean, and the problem is –

CHAIR APOSTOLAKIS: Does anybody model it? I know they look at time. I know that time – I don't know how many things people are looking at. But this is telling me, this is a very important thing.

MR. PARRY: You could argue that the TRCs do it. They are based on simulator data.

CHAIR APOSTOLAKIS: The TRCs is an attempt.

MEMBER STETKAR: If the curves on the TRC is based on simulator data, that's right.

MR. BLEY: It seems to me, anytime if you're doing an HRA, and if the time it might take to carry out the action is on the same order as the time that is really available –

CHAIR APOSTOLAKIS: Then you should worry about it.

MR. BLEY: – then it means that you can shift from success to failure quite easily just by slight things affecting it.

CHAIR APOSTOLAKIS: Absolutely.

MR. BLEY: And you need to observe that.

Now sometimes, if I can just hit on something John said earlier, sometimes it means, maybe you can redefine your HFEs and your PRA a little bit so that that's not – if in fact you built your model a little conservatively or maybe the other way around, maybe you can remodel things a little bit to pull out this time dependence a little, or maybe it's a flip of a coin kind of thing –

CHAIR APOSTOLAKIS: I think what you are doing is, you are addressing the question of how to handle it. And you're saying – and I agree –

MR. BLEY: How one might handle it if they were doing HRA.

CHAIR APOSTOLAKIS: Yes.

MR. BLEY: But if you don't recognize that's the situation –

CHAIR APOSTOLAKIS: That's right, that's what I'm saying, rephrase what I wanted to say.

This is not – this is telling me the time when it's comparable, when it's comparable to the available – should not be treated as just another PSA. That's what this is telling me.

It's not just a PSA. It should be a focus of the analysis. Because this is the very ability that is important.

MEMBER STETKAR: The other part, though, of crew-to-crew variability is, one of the things that probably will eventually come out of this process, one of the things I'd be interested to see is, do you see the same degree of variability over the whole scenario, not to the point where did they successfully equalize pressures before the ruptured steam generator was overfilled?

The time to reduce pressures –

MR. BLEY: Are you asking just to the last one, or all the ones along the way.

MEMBER STETKAR: No, no, after the last one, because that is kind of the integrated task, that is kind of the integrated task. In some sense I don't know. I mean I actually don't know whether the variability that you see within this particularly focused snapshot task is due to consistent crew-to-crew variability over the integrated task, or is it due simply because this is where we sliced our snapshot.

MR. BLEY: If my crew did it two times in a row, I might slip from one side to the other.

MEMBER STETKAR: Well, no, you might not see the same variability over the integrated task, because people might be delaying this – doing this very rapidly, but delaying longer because they focused more on this, or something like that.

I just don't. It's part of this variability – I believe there is actually variability. But part of the range of this variability may be just due to this particular action and the particular slice that you took through the whole integrated crew performance.

So it might be worthwhile kind of following up on that concept, as you finish – as you look at – but when you look at those other actions, also keep the entire scenario perspective.

CHAIR APOSTOLAKIS: I think there are two things. One is what he just said, convince ourselves that this is real, under certain conditions – not always, but sometimes.

And the second question is, if it is real, is anybody willing to change their model to include this reality?

You don't have to answer now. But I see this as part of this suite of models. Under certain conditions, as Denny said, you know, it's clear that the time available is very large. You don't have to worry too much about the variability. In this class of problems, there are comparable times you have to do something about it. You convince yourself that it is important. It's not an artefact of how you model things perhaps.

You look at the whole scenario. This is really the kind of insight that it would be great to derive from this, and there we should be willing, because sometimes people are very defensive defending something they have already done, to modify whatever model is necessary to accommodate for this, in terms of the focus.

MEMBER STETKAR: In terms of the calculator it does that.

CHAIR APOSTOLAKIS: Well, then let's all use the –

(Simultaneous speakers.)

CHAIR APOSTOLAKIS: I'm sorry, but that is the answer if you tell me that.

MR. PARRY: No, all I'm pointing out to you is that there are some people who have thought about these issues. And I don't have Jeff on the phone. But he could tell you that that is the way they would handle it.

MEMBER STETKAR: Well, in terms of this exercise, rather than recommending changes, it's pointing out the fact that a methodology that indeed does either quantitatively in terms of the parameter values that you use or some other way account for the variability in terms of time.

CHAIR APOSTOLAKIS: Well, let me address the question.

The whole idea of the original ACRS letter was let's all agree on an approach. EPRI does it, then NRC should be doing it also using that calculator for the class of problems where it applies.

And this is not the case right now. We are not doing it. That's all I'm saying.

If they have the brilliant idea, and they did something, more power to them.

If we agree that in a certain class of problems, this is the way to go, then I'd like to see the agency here using that model.

Then for another class, if something else that was developed here is better, they should be using that, too. Then we will be in a state of affairs where we will say, yes, we have made progress.

I don't think that people have never thought of these things. All I'm saying is that we have different approaches.

Then if you guys agree that a calculator is one way to do it, or the way to do it, then perhaps you will review it more carefully; you will express some views; and the EPRI guys will come – one of my complaints as you know is that these models have not really been scrutinized the way we scrutinize hydraulic models. Of course it's because these are more accurate than thermohydraulics, right? But the time should come when we should do that; that's all. I'm not claiming that nobody has ever thought about it.

Because there you see – that's another thing – even if you go with the EPRI calculator, they really give you I think one curve, don't they, for this condition.

MS. LOIS: That's the limitation of the ACRS.

CHAIR APOSTOLAKIS: Yes, they define the conditions and then they say, if the time is much longer then you switch to the three, the event three. And so on. But what for example if you want to bring the other performance shaping factors, and according to those you adjust the curve. There are all sorts of things you might want to do.

MR. PARRY: It's not perfect.

CHAIR APOSTOLAKIS: No, it's not.

So what are you telling us now?

MR. MASSAIU: I told you already these ones. So we just keep it. These ones just to show that in the base case also that the failing crews -- they later also overfill the generator.

We have seen these – they are the same. These are just crews that each was selected for – in the analysis.

And just – this one just to point out that so-called failing crew N overfilled the steam generator –

CHAIR APOSTOLAKIS: Which crew was this? M?

MR. MASSAIU: M.

CHAIR APOSTOLAKIS: M?

MR. MASSAIU: Later on in this scenario, when they stop the site injection.

MS. LOIS: Salvatore, I believe that the remaining were kind of backups at the time? Do you want to discuss the remaining of those slides?

CHAIR APOSTOLAKIS: Twelve looks interesting, but I don't know how many more of those you want to do.

In other words, can you look at things very quicky from the 12 through 14, which ones are –

MR. MASSAIU: Yes, I can go very fast on this one.

CHAIR APOSTOLAKIS: You don't have to go fast. As long as you explain what's going on. But if it's repetitive. So pick one.

MR. MASSAIU: This one is the base case.

CHAIR APOSTOLAKIS: This is 12. Okay.

MR. MASSAIU: The base case. And you see that on the top here how many crews did that action. And basically in most cases all crews – or most crews did the same things. So it was quite – pretty well cooperation on what was expected.

CHAIR APOSTOLAKIS: Average crew for base case. Average crew – which one is the average crew?

MR. MASSAIU: It's the --

CHAIR APOSTOLAKIS: Is it --

MR. MASSAIU: Well, in this case you take for example for each individual crew the -- when the manner of the injection. We have checked when they have done that, and we have the time.

What is the average?

MEMBER STETKAR: How large was this break? Was it a single tube break or less? You didn't get automatic safety injection on this.

MR. BLEY: You would have.

MEMBER STETKAR: Eventually. So this was roughly kilograms per second, do you know?

MR. BRAARUD: It's one tube, I think. In the base case it will, if you don't trip manually you will have automatic trip after some minutes.

(Simultaneous speakers.)

MEMBER STETKAR: I was just curious, the timing of the manual reactor trip.

MR. BLEY: I wanted to just – if it's a relatively – anyway, go on.

MS. LOIS: Are you an ACRS technician?

MR. BLEY: I'm not ACRS. I'm an analyst, and I did the ATHEANA analysis. And there is one thing on this one that kind of had me thinking about something that John asked about earlier.

In the plots you gave us, the – whoever was making that run to give us some hydraulic plots of the analysis, manually tripped the reactor in about seven minutes.

And in talking with operators in the U.S.

here, they talked about how – and I understand most of Europe would be the same way – with the low levels of radiation we had in the base case, people would probably take some time trying to figure out where that was coming from, unless they were in a simulator expecting something, and acting in a way different than they would in the power plant.

And everybody except one took this thing in less than a minute on very low levels of radiation that were far below any requirement that would force them to do that.

It got me to thinking about John's question –

(Simultaneous speakers.)

MEMBER STETKAR: – operators are generally reluctant to trip reactors in Europe – I hate to say this but it's true – because availability is a very strongly encouraged parameter in Europe.

So manually tripping a reactor in Europe, especially through reactors that don't trip for years, is a really severe response. I mean that is something that, if you do that and there was not good justification, you might lose your job.

So 13 or 14 crews tripping the reactor very rapidly here –

CHAIR APOSTOLAKIS: One minute.

MEMBER STETKAR: – that's part of – in terms of the timing. I'm not arguing here in terms of the scenarios here for HRA, but in terms of the timing, is this another feature of the fact that these guys knew they were in the simulator scenario, and that they would be expected to perform very positively and very rapidly because that's what they're rewarded for in a simulator environment, perhaps not being rewarded for that in the real world?

MR. BRAARUD: I don't think that is the most important factor in this case based on the observations from the analysis. It looks like the crew, the integrated search, that they have a quite large linkage, so they have quite clear responses to main parameters, that they will actually get through them automatically.

So based on that data it's a manual trip because they are quite sure that it will trip anyway.

MEMBER STETKAR: So you don't think you have as much – okay. I mean you have a lot more experience.

(Simultaneous speakers.)

MEMBER STETKAR: I mean if that is the case, and it was clear that they were headed in the direction for an automatic trip, then that's not necessarily surprising.

It's just the – it's the trajectory that we are not sure about.

MEMBER ABDEL-KHALIK: But if the baseline scenario that was provided, the crew trips the reactor manually after 7-1/2 minutes, and these people trip the reactor after a minute, this is a very very slow scenario.

MR. BLEY: It would have triggered a reactor trip in about 11 minutes.

(Simultaneous speakers.)

MR. BLEY: I'm trying to figure out how we are going to handle modeling these guys in the simulator.

CHAIR APOSTOLAKIS: I think Susan is trying to say. Susan, go ahead.

MS. COOPER: Yes, thank you, George.

I just wanted to amplify what Dennis was saying, as again part of the ATHEANA team.

We actually spent quite a bit of time debating the significance of the manual reactor trip. We could have driven the time trace. And I thought I remembered it – maybe, Dennis, I remember it incorrectly – but I thought that we determined that the manual trip that was given to us in the time trace was just before you would have received an auto; very very close.

MR. BLEY: No.

MS. COOPER: And we were – we were – supposing that, while they wouldn't want to trip the reactor, on the other hand beating auto trip was supposed to be a good thing.

MR. BLEY: That had to do with safety injection as I recall. But reactor trip, they were well ahead of it. Safety injection, they were just before it.

MS. COOPER: No, I'm not saying safety injection; I'm saying reactor trip. They manually trip the reactor just before it would have automatically tripped. We determined that from, I can't remember which parameter, but from looking at the time trace they just jumped in advance of auto trip. The Halden team, whoever was in the simulator tripped it.

So that was part of our rationale for the timing of trip, and that's why the manual – the trip was a force, it was a manual trip, and that of course started whatever time was taken up to that point in time, you weren't in E0, so that time was just time that was lost to go forth and execute the rest of the steps and get into E3.

So yes, I think it is interesting to see that the crews, and the average of all the crew times to manual trip is significantly shorter than that in what we were presented in the time trace.

CHAIR APOSTOLAKIS: Okay, shall we move on?

Salvatore, do you have anything else to day? Your title was results. These were the results you observed; not the results of the analysis.

MS. LOIS: No, that's going to be covered by John.

CHAIR APOSTOLAKIS: Okay, okay. So I find 18 interesting, unless you have something else before it.

MR. MASSAIU: If you want to connect to this latest comment, I think, we estimated that automatically in the base case would be about four or five minutes after the start of the leakage, so in this case it would be some minutes – the manual trip would be some minutes before the automatic.

MEMBER ABDEL-KHALIK: That's not –

MR. BLEY: Automatically they come at eleven -- they --

CHAIR APOSTOLAKIS: They knew that there

was going to be an automatic trip so they just took action; that's what it said.

MR. BRAARUD: Yes, one example is page 159. That's a transcript for one group. We didn't want to go into details. But you can look.

MR. MASSAIU: If we put the two scenarios side by side, we can see that there are – well, the timeline splits, but if you read the time from the timeline, you will see that there are quite significant differences in how long a time it took for transferring to E-3. It took three times longer to do the transfer in the complex case.

One, the time used for isolating was just one minute longer in that complex case. So it was more or less the same time to isolate, but identification was the difference.

CHAIR APOSTOLAKIS: In the table you showed earlier there were some crews that took a very long time for some things. If I go to the time diagram on here, would I be able to see that somewhere?

MR. MASSAIU: Not really, because these are averaged over all crews.

CHAIR APOSTOLAKIS: That was behind my questions.

MR. MASSAIU: And I look to step 21,

because many crews diverged from that step, and so the – a large amount of variability comes from what happened at that stage.

The next slide is about that stage, and I call it a crucial decision point, certainly meaning that you have different ways of proceeding crossing at this point.

The difficult thing about this point was that they had to assess the RCS pressure, and this assessment was not that easy to do at that point. It was just started normally, it was just starting to –

CHAIR APOSTOLAKIS: This is the complex scenario?

MR. MASSAIU: Yes.

And of course just starting to train at that point, and that is also something that creates difficulties for the crew to use time to assess parameter changes over time.

At least this is what we observed. They had some problem doing that. At the same time they are to consider conflicting information at this stage, because of course although many noticed that there was a suspicious level in SG1, there could have been an alternative explanation for that one. So we have crews that are controlling all the status of the second line system, all the feedwater, auxiliary feedwater, and there might be a leakage from that side into the steam generator.

Some crews, as we discussed, were also thinking about what's happening in the turbine hall, because they knew that something had happened there.

So it was not that easy for the crews to assess, and so they took different decisions. And some crews made the transfer to E-3 directly from there. However, there is not an explicit transfer point. And the four crews out of 14 – other crews, they went to the procedure ES-1.1, which is safe injection termination. And they terminated the safe injection. They had to restart, so they went back to E-0, after that they looked there, they normally -- they transferred from a step #19 those crews, although they deterred radiation again. But they made the analysis assessment anyway.

Two crews, they used the foldout in ES-1.1, and so an interesting thing is that although you have a transfer point, it is not sure that you are going to use that as you see in these results.

Two crews used the transition from steps 24, 25, which doesn't necessarily mean that they followed those conditions for transfer which we have in those statements. But they transferred while they were working on those steps.

And then you had also two other ways of transferring. The last one is a crew which manually tripped the reactor, so they had to manually isolate the safety injection into the steam line.

MEMBER STETKAR: This is in the complex case, they took the reactor and manually isolated the steam line before?

MR. MASSAIU: Yes, and then they probably got some radiation too.

MEMBER STETKAR: That's an incredible response. And I don't – I mean it both in terms of really, really fast, and I doubt that anybody would behave that way in the real world. That's really interesting.

MEMBER ABDEL-KHALIK: That is less than 30 seconds.

MEMBER STETKAR: That's really interesting.

MR. MASSAIU: Yes, a few reactors hit the manual reactor, it was very, very fast.

MEMBER STETKAR: In the complex case?

MR. MASSAIU: Yes.

MEMBER STETKAR: After they operated in the plant for several years without even having a reactor trip.

MEMBER ABDEL-KHALIK: Were they standing there looking at RCS pressure, and they saw it and they just tripped the reactor, or what? I mean this isn't realistic.

MR. BRAARUD: They were reacting rather fast. You can see them quite good on the large screen. And they take a quick decision.

MEMBER STETKAR: No, that's simulator gamesmanship.

CHAIR APOSTOLAKIS: That's what?

MEMBER STETKAR: Simulator gamesmanship.

MR. BRAARUD: But it was only two crews, wasn't it? That's all.

MEMBER STETKAR: Yes, it was only two crews, 14 percent.

MR. MASSAIU: So I have some more details. I don't think I need to go into those.

CHAIR APOSTOLAKIS: This is 18, 18, I thought was interesting. Maybe 17 is the same – no, 18. Oh, your 18 is different from my 18. My 18 is your 17.

Yes, the – what message do you draw from the right-hand side column? I mean why did you bother to identify them as knowledge based or something else?

I mean do you plan to learn something from it?

MR. MASSAIU: I think this was a point which created some difficulties from the analysts' side, because many were assuming that the procedure transfer would have been out of the transfer points in E-0, and so there was a great deal of discussion about how they did the transfer, and what was the reasons for transferring to the E-3 procedure.

CHAIR APOSTOLAKIS: The thing that confuses me a little bit, and then you can address this, you said earlier – I mean the group said – that there is a procedure for the complex scenario.

MR. MASSAIU: It's the same procedure for the two scenarios.

MEMBER STETKAR: There is only one procedure.

CHAIR APOSTOLAKIS: One procedure, okay, and it includes though the complications.

MEMBER STETKAR: Theoretically it includes everything.

MR. BLEY: It does the diagnosis; that's what it's for.

CHAIR APOSTOLAKIS: But isn't the idea behind having a procedure to avoid finding the – putting the operators in a knowledge-based situation?

MEMBER STETKAR: But you can't write a procedure for every foreseeable situation.

CHAIR APOSTOLAKIS: Not for everything. But my goodness, look at all these knowledge based –

MEMBER STETKAR: Well, that's true. But that is what happens in the real world. That's real – from a – I found that really interesting and really useful information, especially if you are evaluating – again, the purpose of this exercise is to evaluate different HRA methods and the applications of those methods.

And some of the older methods that force you to assume either skill, rule, or knowledge-based behavior, and to particularly focus on rule based behavior in this type of scenario, this is really interesting information.

CHAIR APOSTOLAKIS: From what point of view is it interesting?

MEMBER STETKAR: The variability that you get in the application of a methodology, or the method to assess this type of situation. Because this is real.

CHAIR APOSTOLAKIS: After TMI there was a strong message sent by the industry or the community that we will really go to procedures, rule-based behavior, to avoid – or to minimize the knowledge based actions.

And I look at this table, and I'm scratching my head now.

MEMBER STETKAR: Well, George, if you read what's in the report, you had people who figured out it was a tube rupture and couldn't find a way to get there, is mostly what it red lighted to me. And a couple of them found the fold out page that gives you a reason to go there, and some others didn't.

And they just couldn't quite find their way through the procedure. A couple started the procedure – several started the procedure over again and said, we must have missed something.

MR. DANG: A side finding of this particular experiment is that this procedure step, one of the transfers, is flawed. And we had the comment that in other plants that procedure step, that specific step where they had trouble satisfying the criteria, has been changed, and has additional criteria which would prevent them from needing to go to the knowledge base.

CHAIR APOSTOLAKIS: I think that makes more sense.

MR. BLEY: But if they had stepwise kept

going, they would have gotten to a step that would have carried them over that. It was when they thought they ought to be going and couldn't – at least that's what I read from the words that I think you took from them, that they decided they had to do something.

MEMBER ABDEL-KHALIK: To me this knowledge based thing is for early action rather than where the procedures would have gotten them there.

MR. FORESTER: As they said, they sort of diagnosed the STTR so they were looking for a way to get there. And some of the steps that would've gotten there required sort of an assessment over time to validate what was going on.

So it was like they were looking for a quicker way to get there, because they knew that the situation – even though some of these other steps might have gotten them there eventually if they had been more patient about it – but they knew what they had, so they decided to take care of it.

MS. COOPER: Susan Cooper. I just wanted to add, when the ATHEANA team was looking at – starting to look at this complex event, the – we had three former operators who are now NRC employees who were helping us out. And one of them did a search for events in the U.S. And he also came up with some information notices as well.

Through that information as well as his own experience being a former operator of a Westinghouse plant, he pointed out that there are a number of differences or changes that have been made to procedures here in the U.S. to address some of the problems that we're seeing in the simulator, like the fact that the radiation indications weren't present at the time when they reached the steps that would have asked that question, or that there might not be any at all.

So there have been some modifications to U.S. procedures really in response to previous incidents where people have gotten sort of tied up in the procedure and not getting where they wanted to go.

MR. BLEY: Said, there is a report that is done for the NRC, it's a NUREG/CR. It was done by the group that used to be at Westinghouse, Dave Woods, Emily Roth, Randy Mamaw. I forget exactly when it was done; it was about 10 years ago. And they had come up with a situation to identify very cognitively difficult situations. And after they had done that they ran a bunch of U.S. operators through a Westinghouse simulator to see how they did in these tough scenarios, and almost every crew started getting out of the procedures or jumping through them because they weren't essentially getting satisfaction. They weren't thinking they were going to solve the problem before they had real trouble.

It was a real interesting report, and it was funded here out of NRC.

CHAIR APOSTOLAKIS: Okay, are you done?

MR. BLEY: This is a tough scenario.

CHAIR APOSTOLAKIS: Are you done?

MR. MASSAIU: I am done. The isolation was simulated under two conditions.

CHAIR APOSTOLAKIS: Yes, but I think we got the message of what we are doing.

Very good. Thank you very much.

Who's next?

MS. LOIS: Per Ovind again.

CHAIR APOSTOLAKIS: He's not on the list, is he?

MS. LOIS: He is.

CHAIR APOSTOLAKIS: Where?

MS. LOIS: He's item #5.

CHAIR APOSTOLAKIS: Yes, you're right.

(Simultaneous speakers.)

MR. DANG: I have about an hour and a half reserved, and I think we will manage it.

CHAIR APOSTOLAKIS: You guys are leaving at 4:00.

MR. MASSAIU: We leaving at 3:00.

CHAIR APOSTOLAKIS: Oh, I thought it was 4:00.

MR. FORESTER: We're staying as long as you want.

(Off the record comments.)

MS. LOIS: So this is the package now, examples?

CHAIR APOSTOLAKIS: So it's not Forrester?

MS. LOIS: No, it's Halden results.

CHAIR APOSTOLAKIS: Right.

MS. LOIS: So we are not doing the presentation we were supposed to.

CHAIR APOSTOLAKIS: Why not?

MS. LOIS: Well, actually, John should have explained a little bit about the HRA methodology before we explain Halden, but that's fine.

MR. FORESTER: That's fine. It works fine this way.

CHAIR APOSTOLAKIS: But we are skipping you completely?

MR. FORESTER: No, no, I'll be back. As soon as Salvatore, then Olvind, then we'll move to slot six.

PARTICIPANT: Slot four.

MR. FORESTER: Slot four, right. I'll do slot four, and then I'll do six back to back. Four and six will be back to back.

CHAIR APOSTOLAKIS: Whatever you want to do.

MR. DANG: The reason we moved five was to accommodate their travel plans, was the reason we moved five up.

CHAIR APOSTOLAKIS: I know.

MR. FORESTER: But it works just as well this way, maybe better.

HALDEN DATA COLLECTION, ANALYSIS AND EMPIRICAL RESULTS

MR. BRAARUD: Okay?

Now I'm going to, as we saw on the previous presentation, the material from the analysis. We saw a very rich quantity of material about, also about why did the crews use a long or a short time.

So we had an overall picture in the first presentation. And in this presentation, I am going to give some examples, brief examples, not very thoroughly, not vary over all results, although some of the explanations for the performance.

Vinh will follow up with the whole – the total analysis. It's summarized into performance shaping factors.

Okay, so I will quickly give a few examples on three selected PSFs.

So if you take the base, STTR, as an example. The – a quick overall summary of the quality analysis is that the crews, they performed the scenario with a good understanding of what is going on.

Early in the scenario, they detect important alarms. They interpret that this is likely STTR. They have activity alarms. And they quite quickly, this is an example, trip the reactor based on the plant simulator. Plus an indication that they are quite sure about what is happening in this scenario.

And they enter the diagnosis procedure, E-0, and the observations point to that by running through this procedure they don't find any additional problems. Actually the procedure confirmed what they were expecting from the beginning of the scenario.

And they find the activity indications in the STTR step, and they seem quite sure that they actually have an STTR, and transfer to that procedure.

One issue they have quite good training on the base type of the STTR. It follows a training program at the plant. It's often included in every year's training, the base type of the STTR, in some form.

So this is the scenario. They were not very good. And also in the interior, some examples of comments is that this was a standard tube rupture. It's not difficult. We have run this often. It's very similar to the ones we had on in the training simulator.

One other PSF or factor is teamwork and how the crew actually organized their work.

In the base scenario, it seems like the crews are not very much challenged on how they actually perform the management on the team. They basically followed the procedure. The procedure guidance is very strong. It's a very familiar scenario.

Yet some variation in the whole, for example, a supervisor managed the team. But this doesn't have a strong effect on the time they actually need to proceed through the scenario.

CHAIR APOSTOLAKIS: All this seems to go against the actual finding. The time – are you talking about the HFE 1A, right.

MR. BRAARUD: This is the base scenario.

CHAIR APOSTOLAKIS: The base scenario? But they are trying to isolate the significant variability. If they have trained, if they think it's trivial, if we've done it before, why did one crew do it in 10 minutes, and the other in 22? Double?

MR. BRAARUD: Okay, there are some variations. That is also the other conclusion. So that is a discussion of how much variance is actually much variance in this case.

It's also related to the performance criteria we set up. But they don't – the variation does not actually lead to that they pass this performance criteria we set up.

MEMBER STETKAR: The difference between one and five minutes would be a factor of five, but the difference doesn't make any difference, so you have to be a little bit careful looking at that.

CHAIR APOSTOLAKIS: No, but all these comments, we've done this often, it wasn't difficult at all, it was easy – one crew took double the time of the other crew.

MR. BLEY: You are not supposed to jump in and solve it. You are supposed to stepwise go through the diagnostic procedure just in case there is something more than you are thinking about. So you are supposed to do that systematically.

MEMBER STETKAR: Maybe it didn't make any difference if they took 35 minutes.

CHAIR APOSTOLAKIS: No, the limit was 20 minutes.

MEMBER STETKAR: But that is an artificially imposed limit. They knew that.

CHAIR APOSTOLAKIS: No, they didn't know that. They

MEMBER STETKAR: They didn't know that they had only 25 minutes. They knew that they had a tube rupture, period. I mean shortly after they had a tube rupture, they knew they had a tube rupture.

MR. PARRY: There is also a significant difference in the times that they actually tripped the reactor, which is where they would start going through E-0. That is a factor that plays into the time to isolate.

CHAIR APOSTOLAKIS: There is some anomaly here, because even if we take what Dennis just said, that means the 10-minute guys did not go through the procedure step-by-step methodologically.

MR. PARRY: They may not have tripped the reactor quicker.

MR. BRAARUD: There is some variation between the crews. There is for example variation in the speed they run through the procedure. Some read the procedures more thoroughly, more slowly, assuring that they have good communication with the assistant reactor operator, while some read quicker through the procedure. That is one factor.

Also the crews vary if they perform a meeting or a consultation or not before they transfer to the STTR procedure.

Okay, so I can moderate this a bit. There is some variation. But not as strong in the complex scenario.

MEMBER STETKAR: I was going to ask you when you get to a complex scenario, but it's probably worthwhile interrupting here, you mentioned here that the crews had an expectation of what the scenario was, and that the procedures confirmed that expectation.

Is that something when you are doing the results, you are going through the DVDs and actually watching them perform, is that something that you saw across all the teams, that they made a diagnosis and then basically confirmed that through the procedures? Or did they use the procedures to finally reach the diagnosis?

Do you understand what I'm asking?

MR. MASSAIU: You have both cases. You have both cases. If you remember there was a very slow crew in the complex case. That crew was hardly trying to follow the procedure. They had a poor understanding of what was going on. And that was probably the reason why they were trying to stick to the procedure, which took them a long time.

And the transfer – the ground for transfer was radiation, because they got radiation after some time, some radiation.

So you remember a situation in which they know what – or they suspect strongly they have a rupture, and they are trying to find confirmation in the procedure. But some other –

MEMBER STETKAR: It's just probably more relevant when you get to the complex case because of this knowledge base transfer to E-3.

MEMBER ABDEL-KHALIK: Was there any observed trend as to the relative level of experience of the supervisors versus the operators, and whether that relative experience level would have an impact on the response?

MR. BRAARUD: We haven't analyzed this in detail yet, but general observations could point to supervisors with very low experience performed differently than those with over a minimum experience. But that is some of the observations we haven't analyzed in detail yet.

There might be some differences according to the experience in the position.

Okay. So if you go to the complex scenario. There is quite a difference if one of the structures is actually the complexity, how difficult is it to understand scenario.

And in this case they have a steam line break in the first part of the scenario, and the plant signature actually is similar to a steam line break.

Several crews actually commented in the beginning that this looks like a steam line break, which that is a correct assessment in this situation but will not necessarily help them through this scenario.

There was some thought, maybe this is a steam line break, and they commented it briefly, and started the diagnosis procedure.

Later in the scenario there were some examples that the crew interpreted the process development according to their initial hypothesis. For example when they interpret the imbalance between the SG levels, the one that ruptured and the other ones, some crews actually interpreted this as support for the steam line break hypothesis.

This also has to do with the control of feedwater manually, so they were not quite sure what was the effect on this steam level, was it something leaking? Was there bad control? Or what was actually the situation.

Also the cool down that followed from the break confused some of the crews. How should they actually interpret the RCS pressure? That could be an indication of the STTR, but they interpreted it as a result from actually the initial event.

So examples from the DVD for the analysis. Reactor operator suspects this is a secondary break.

One example of the interpretation of the RCS pressure just from the analysis. And also they don't find the expected T indication when they pass the step. There is actually no support there.

So actually in principle all crews process this step, since there is nothing about steam levels in that step, except for one crew which was aware, as Salvatore mentioned, was actually aware during the scenario.

MEMBER STETKAR: I hate to keep asking, but more questions come up.

Does this plant have high pressure high head safety injection pumps? Or what's the shutoff head on the high pressure injection pumps on this plant?

I don't know if it makes much difference, but I was just curious. Is it 110 bar, or is it 170 bar? Do you know?

MR. BRAARUD: I don't think I could tell you.

(Simultaneous speakers.)

MR. BLEY: – might have given us a clue to that, but I'm not sure.

(Simultaneous speakers.)

MEMBER STETKAR: I'm not sure. It just popped in when we were talking about pressures and things like that.

MR. BRAARUD: There are some examples also this missing activity indication. It's actually delaying crew, or hindering them in making the actual STTR diagnosis.

MEMBER ABDEL-KHALIK: Just out of curiosity, what did the automatic trip happen on?

MR. BRAARUD: Likely it would be the pressure or level. Pressure, I guess.

MEMBER STETKAR: Well, it could be high steam flow too. It depends on how big the break is.

MEMBER STETKAR: It could have been. You're right. I thought I remembered that, but I'd have to go back.

MR. BRAARUD: So other examples that the crew actually discussed, they think they have an STTR, or discuss also the secondary break. And they search for these activity indications, but they don't find them. And actually continue in the procedures.

While as Salvatore explained when they are actually not able to find a good way through the procedure, they actually search for a transfer point, or in some cases actually make what we call a knowledge based transfer.

But this is some of the basis for the complex scenario some were given, was assessed as complex on this PSF rating.

As far as training and experience. The actual crews, they had previously for many years ago trained on STTR scenarios without radiation, but not this scenario, other plant base.

And in the interview, some operators pointed to the fact that they knew the different parts from before, STTR, and such, but this combination was not familiar to them, and that made it difficult.

What we actually found for the complex scenario was that this challenge, the team dynamics and the various processes. Since procedures do not guide them directly, it's much more up to them how they organize, their detections, how they bring it together, discuss it, and come up with a good plan for the scenario.

So these results also point a bit to the direction within performance shaping factors; that when you don't have your specific items and scenarios are unfamiliar, the team management, the supervisor, becomes important.

So we observed the variability in how well the supervisor was able to establish a, we can say, a good interpretation and planning process in the crew.

Typically some crews continued to work while discussing. They detected all important pieces, SF level for example. But they weren't able to step back a little, discuss what is going on, and make the correct plan.

MEMBER STETKAR: Out of curiosity, you make the observation that the supervisor is the key. Did you in general, did you feel that the supervisor at least in these 14 crews was acting as more of an integral part of the crew, or as a rather more independent orchestrator if you will?

Were there any general conclusions that you could draw from that? In other words, you had a three-person crew but a supervisor and two reactor operators, let's say. Were they operating together as a three-person crew, or were they operating as a supervisor and two reactor operators?

MR. BRAARUD: Well, the main finding is that there was viability in how the supervisor performed their work. But generally the supervisor has a more overview function than the reactor operator and the assistant reactor operator.

MEMBER STETKAR: And that's what you observed in practice? Or that's the way it's supposed to be?

MR. BRAARUD: That is the way it's supposed to be. But in some cases the supervisor was more actively involved in the process operation. Like if the crew had a problem the supervisor gets somehow a little bit driven into the situation, starting to interpret and suggesting how to interpret the procedure steps, and in some cases actually forgot a little bit the team management part, stepping back and having the overview, and calling together consultation within the crew.

MEMBER STETKAR: I think that's an important finding. I thought I read that, and I think that's an important finding, because there are so many not necessarily the methods but the implementation of the method that treats the supervisor as the independent eyes and ears who will eventually solve the problem even if the reactor operators are too enmeshed in the details.

So that your observations regarding the relative independence of the supervisor I think is an important point.

MR. BRAARUD: That varies; that's an important point.

MEMBER STETKAR: And the variability in that, that you can't – it's not correct to assume that it's an integrated X-person team, and it's not correct to assume that the supervisor is independent.

MEMBER ABDEL-KHALIK: So if you had a fourth person as an STA for example, these results would be totally different?

MEMBER STETKAR: It's not clear; they didn't have a fourth person.

MEMBER ABDEL-KHALIK: If there were a fourth person.

MEMBER STETKAR: It's not clear. They might have got their act together, or they –

MR. BRAARUD: Okay. The rest of the presentation like I said is some more details on how the crew interacted, but depending on time I can go through it. I think the main message is that there is viability and there is some dimension describing how the supervisor managed the team, and how they performed the consultations within the crew.

So that is some of the explanation for the long performance times and the short performance times in the complex scenario, the team dynamics and the management.

Just to sum up, it's actually the base scenario, it's not difficult, relatively. They work in a quite similar way; even there is some variability as we discussed.

While the complex scenario creates a challenge to the group, and one important part is actually the team dynamics supervisor. Although there is interaction between the type of scenario and team dynamics.

MEMBER STETKAR: I'm going to follow up on something Said brought up before; and that is, this is even more interesting in the sense of tracking through all of the actions in both the base case and the complex scenarios to see whether crews in a sense are correlated. The crews with the best team dynamics and the best interactions, do they perform – I don't care whether they perform better or worse, but do they at least perform consistently, so that indeed there is a – that would tend to reinforce this variability among crews regardless of scenarios.

MR. BLEY: That seemed to be a result in the report was that the only – as I recall – the only positive factors were the three or so crews that had – were real good on those two aspects.

MEMBER STETKAR: In the report, though, everything is focused around the analysis of the base case and the analysis of the complex case. I think what Said was asking for was, is crew A correlated through both the base case and the complex case of this HFE for the SGTR and all other HFEs for the SGTR? That would tend to reinforce this variability on a crew-to-crew basis consistently, regardless of the action, perhaps regardless of the type of event.

And that would tend to much more strongly reinforce this – the need to account for that.

MR. BRAARUD: Very interesting, and we are in the process of doing that.

That concludes my presentation.

(Off the record comments.)

CHAIR APOSTOLAKIS: Okay, Vinh.

MR. DANG: Okay.

My presentation covers the last step of the derivation of the reference data, the empirical data.

Basically what Salvatore and Per Olvind have presented to you is what we've seen in the simulator; what each of the crews did; and this step is the step where we identify what the main performance issues are for this HFE, when we look at all the crews.

Our correspondence to this last step down here of coming up with the experiment outcomes.

I'll run through this session. I'll go through it quickly. It's to define for you very briefly where the driving factor is. I think you know what it is, but I will just make a couple precision points there.

Discuss the process of taking these 14 observed performances for each scenario, and coming up with the factor ratings, or the identification of the main issues for that.

And then I'll summarize what these various factors are.

Okay, so basically, a driving factor is a PSF. We take them both positively and negatively. We identify those that are positive as well as those that are negative.

And to be a driving factor this PSF has to be important, which is a combination of the weight, meaning the performance level, the difficulty is not indifferent to this factor; and it has a rating. It's either positive or negative, this driving factor, as opposed to being nominal.

We use a set of performance shaping factors. These are I think the usual suspects if you will.

The next shows you the correspondence of the ones we chose versus the HRA good practices.

CHAIR APOSTOLAKIS: There's nothing here.

MR. DANG: That's odd, isn't it?

(Off the record comments.)

CHAIR APOSTOLAKIS: Okay, thank you.

MR. DANG: And you thought I was done?

CHAIR APOSTOLAKIS: I thought you were done.

(Laughter.)

CHAIR APOSTOLAKIS: What do we do, then?

Do we share? I guess we share. Do you have any more? Oh, okay.

Anybody else? Okay.

MR. DANG: Sorry about that mixup. Okay.

So this sheet shows you for your perusal maybe later the correspondence between the ones in the good practices as well as the HRA PSFs which are more intended for retrospective use at this time.

In terms of the factor rating scale, we basically used a seven-point rating scale going from very good to very poor, or very low to very high.

This qualitative rating of whether it's somewhat poor or poor or very poor is useful for us to remember which is the most important. But in the comparison which we will discuss later, we don't use it so strongly at this time.

CHAIR APOSTOLAKIS: So what does poor mean? Very negative?

MR. DANG: Yes.

CHAIR APOSTOLAKIS: Very poor under negative means a very strong negative influence.

MR. DANG: Exactly.

CHAIR APOSTOLAKIS: Why didn't you say strong then?

MR. DANG: It could be very strong,

depending on the –

CHAIR APOSTOLAKIS: Or very high, you have underneath very high.

MR. DANG: Or very high, yes. It really depends on the actual factor. Okay?

Okay, this picture you have seen.

CHAIR APOSTOLAKIS: And who did this rating? Who evaluated this?

MR. DANG: We used the qualitative information that has been presented to you, and correlated it with what happened. I mean what did they do to determine which ones were the strongest. It's done by committee.

CHAIR APOSTOLAKIS: A committee of people who are running the shops?

MR. DANG: Yes. I mean in practical terms I drafted the first version, and a lot of people commented, and we eventually converged on which ones should be strong or very poor; which ones should be poor; and so on. And we double checked against beta to make sure that it was reflective of the qualitative data that underlies these ratings.

CHAIR APOSTOLAKIS: Okay, so this is a consensus then?

MR. DANG: It's a consensus.

So my presentation deals with this last step on the right, which is to take the 14X2 analyses per crew, and to aggregate it over all the crews to come up with the main operational expressions, and the driving factors, with their ratings, which are then used in the comparison.

Then the next steps are the procedures, the procedure by which we think this is true should come from 14 to one set of issues. We basically looked – rather than looking at all of the 14 performances, we tried to look at the best and the worst, the two ends of the spectrum, to be able to sort of contrast really what – if they're really bad what seems to be driving their performance. And if they do okay, or they do well, then what seems to be driving their performance? And to try to see that contrast rather than the strict continuum where you would just see where they did a little bit worse and so on, where you would have a very hard time to see what is driving the difference. To see the most difference we looked at the two ends.

And to make sure that we do not key only on one best or one worst, which could be idiosyncratic, we picked two or three of them.

CHAIR APOSTOLAKIS: So what you are going to decide then, what you are going to conclude, is that the PSFX could have a very negative influence on performance. It could, because you are looking at the worst. That doesn't mean it will.

MR. DANG: That's correct, yes.

CHAIR APOSTOLAKIS: And vice versa with the good influence.

MR. DANG: Right. Right.

So our first step is to take the qualitative information from all the crews – in this case it's from the best crews and the worst crews, and to put them into this matrix, where you get an overview of well, those are the factors for each of these sets of crews. That comes from the performance data that has been discussed up to now.

The next step is to look at the negative factors in the worst performance, and the negative factors in the best performers. And that tells you really something about what seems to be generally negative in the scenario, but what seems to be driving the worst performance.

So the trend towards worse performance will be in only those factors that you see in the bottom right, since the ones in the bottom left were negative, but those crews seem to do okay.

Then the next step, step three, is to compare up and down among the best performers. They had negative factors. What were the positive factors that seemed to compensate for these negative factors?

And then you can also do of course a comparison of the positive factors among the best and the worst and so on. But really the key parts of are the two where we compare only the negative factors among the good and the bad, and step three where we consider positive-negative among the best, to see how these could be interacting.

Finally, if you had one of these tables filled out for the base case and the complex case, we overlay them because by comparing the base and the complex we may see things that we would – that are there by omission. So maybe it wasn't so obvious in the base case that something was positive. But by comparing it to the complex case where it was negative you see that since it's not negative, then it might have been the positive driving factor.

Now I give you a quick look at the complex case and some of the things that were filled in. On the top of the positive factors, the best performances seemed to have good team dynamics. Indications seemed to be the alternative indications to radiation were available, and appear to have been used by both the good and the bad crews.

Among the negative factors, the work processes seem to be making somewhat of a difference, and then a number of negative factors –

CHAIR APOSTOLAKIS: What is the definition of work process?

MR. DANG: Work processes include the –

MR. FORESTER: I think generally part of that terminology comes from SPAR-H, and I think they talk a little bit about it in terms of how to implement procedures. It has to do with the crew dynamics, the team dynamics, how they interact with one another. It's how they do things in a good systematic way. That's the general – it can be – different words can be assigned to describe that.

MR. DANG: The effective use of meetings would be something that would have said that's good work processes. The consistent use of repeat back and so on. These turn out to be among the good work processes.

I don't know if anyone can add anything?

MR. BRAARUD: No, that's true, and the example of team dynamics is more the management of the team, so the supervisor's role is into that team dynamics. But they are related.

CHAIR APOSTOLAKIS: Because usually work process is something broader, when they want to achieve something to do something in the plant, there is a certain process involving different people, you know, did you prepare the package. Somebody checks it, then somebody else implements it.

That's not what you mean here. You mean the modus operandi, that's really what you mean. Right? The way they function.

MR. DANG: Yes, right.

MEMBER ABDEL-KHALIK: Was all the data collected during day shift?

CHAIR APOSTOLAKIS: Yes.

MEMBER ABDEL-KHALIK: Would it have made any difference if some of the data was collected on the back shift?

MR. BRAARUD: It was all run on the data in this case. But of course day and night could have some implication, but we don't have any data resource in this study on that.

MR. DANG: Okay, continuing with the negative factors, well, the complexity because of the masking, the training in particular the lack of training on the handling steam generator tube rupture without radiation indications seemed to be – while they were trained on it but it did not seem to help. And the procedure guidance which you have heard enough about.

I just want to point out in summarizing the driving factors for this case, we ended up putting practically every driving factor in the negative category. So even though there were positive things up here, they did not seem to drive the overall performance.

The other thing I should mention is what has been highlighted in yellow here, the indications. Do you see what indications – well, on the one hand on the positive side, they were alternative indications to radiation. They were in fact used. They supported this knowledge based response.

On the other hand the indications were poor in the sense that those that were relied upon by the procedure were lacking. We ended up rating the overall indication of plant conditions to be somewhat poor to poor, and we had a special mark here, the asterisks with the plus, which means that there was a positive aspect to this, but it was the minor – like the dominant mode is negative, and there is a minor mode of being positive.

The same thing with the training. We would say it's moderately poor. They do not seem to have enough training on a steam line rupture without radiation. On the other hand they did manage to get through the scenario mostly, so the training is not only poor, and that's why there is also an asterisk across there.

Okay, this is the base case. I don't think I'm going to go into this one. A few more positive factors obviously. And that's the roll up of the driving factor identification process.

Now we will come to the comparison later, but I just want to remind you that these – the fact that these driving factors are said to be positive or negative, how poor and so on, for the base and the complex case, is not the only reference data we are using. Through this we hook all the operational expressions that we found. So that means that if you say procedural guidance is good, we have underlying that why we think this, and if we say procedure guidance is poor in the complex case, we – this – the specific issues in the procedure that actually underlie this core procedure guidance.

And then we use that also in the comparison. So the driving factor identification is sort of the assessment of these factors, but hooked on this are all the qualitative information as well.

And that concludes my presentation.

CHAIR APOSTOLAKIS: So there is time for a break now, I think.

MS. LOIS: I think this is a good breaking time.

CHAIR APOSTOLAKIS: Back at 2:55.

MEMBER STETKAR: What time do you have to leave for the airport?

MS. LOIS: I believe because of the weather prediction and the potentially icy weather condition, they are going to leave right now.

MEMBER STETKAR: The question is, before we break, is there anything else that we'd like to get feedback from them if they're disappearing.

CHAIR APOSTOLAKIS: I don't know. Any questions for them?

MEMBER STETKAR: I've beat them up enough. I was curious whether they had any feedback.

CHAIR APOSTOLAKIS: Any parting words of wisdom? Remember, they have to be words of wisdom.

Thank you, gentlemen –

MEMBER ABDEL-KHALIK: I'd like to follow up on the issue of their willingness to run some of the same experiments at a U.S. simulator. What is the scope of that effort? What would the scope of that effort be?

MR. BRAARUD: One scope of that effort would be to see how generalizable the results from HAMMLAB – for example, running a scenario very similar to those we have run in our lab. For example in a training simulator in the U.S. to see that we observe similar performance times, or we will investigate how similar are the times. Do they use the procedures in a similar way? How is the work actually organized? Is it so that these are similar factors explaining performance, or different factors.

MEMBER ABDEL-KHALIK: Would you have the same data collection capabilities? Would you somehow implement or modify the simulator to provide for comparable data collection capabilities?

MR. BRAARUD: It depends on the status of the simulator. But we would surely like to have good logging possibilities. That is often not that good in training simulators.

But also we would have some observation possibilities, and do some scenario analysis for the actual plant to collect the relevant observations.

So I can't say directly how much work we have to do with that.

CHAIR APOSTOLAKIS: Does the Agency have a simulator? What do we have down in Tennessee?

MS. LOIS: Yes, we have a simulator in Tennessee.

CHAIR APOSTOLAKIS: Would that simulator be used?

MS. LOIS: It could be, but the creation – the better question is, we would like to have real crews, and the simulator facility at the NRC is – instructors that haven't been operating a real plant for quite some time.

CHAIR APOSTOLAKIS: No. But it might be easier to get American crews if you asked them to travel to Tennessee rather than Norway.

(Laughter.)

(Off the record comments.)

CHAIR APOSTOLAKIS: Anything else? Well, thank you very much. Have a nice trip back home.

(Whereupon at 2:41 p.m. the proceeding in the above-entitled matter went off the record to return on the record at 3:03 p.m.)

EXAMPLES OF COMPARING HRA RESULTS TO HALDEN DATA

MR. FORESTER: Okay, all right, well, as

you know, the HRA teams provided their predictions in a format that we asked them to provide it in, it was a Form A and a Form B and HRA documentation.

And then the HRA analysis team summarized those results for comparison with the actual data.

So what I'm going to do now is tell you why we did that, and a little bit about how, and also provide you an example.

So why were the summaries needed? Well, each of the HRA methods attempt to capture the factors that affect performance. And they use that information then to derive the human error probabilities. So that is the information we wanted to capture.

What information are you using to drive those human error probabilities, and what do they identify as the important factors?

But there was some variability in how the HRE teams interpreted what was supposed to be – go into Form A in terms of the level of discussion. And then we wanted to make sure we had a fair representation of what each of the methods have done for comparison with the crew data.

So we did this. There were two assessors that did each of the methods, summarized the results of each of the methods. The assessment team had not seen the Halden results when they did this. And again, we mentioned this this morning, once we had done those we sent them back to the HRA teams, and they could make comments. And we wanted to get some consensus on the representation.

MEMBER STETKAR: The same assessors assess each of the – all of the HRAs?

MR. FORESTER: No, it was distributed. Several of us did multiple, and others might have only done one. But no, it was distributed across the team.

Okay, so the general approach then for summarizing the methods. The HRA Form A serves as the primary basis, and there were two things there that we looked at as we talked about this morning. It was the predictions and the factors driving performance, that is, as specified by the HRA method. And then there was the qualitative assessment of what we've been calling the operational expression.

So we asked for each of those. And we tried to summarize both of those things for the HRA teams, and we wanted to reflect each team's understandings and expectations for the scenario, and then to use that information to support what they identified.

Again we relied mainly on what the HRA teams had provided in Form A. But we did supplement that information with what they provided in their HRA documentation.

Most of the teams did a really nice job of actually documenting the HRA results. As you would have for a PRA and the IPEs. So we wanted to use that information to make sure we were getting all the information in that.

And then to the extent needed we'd also use the information from Form B which was the H2ERA taxonomy to help clarify possibly some of the things they had said.

When we did the analysis – I think we quickly covered this this morning – we looked for both the positive and the neutral negative, and the positive or neutral influences, and we broke this out in terms of what was influencing the diagnosis, or the cognitive part of the action of the human failure event, and the response execution part.

And then again we used information from any other sources available to help us interpret what they had identified as the driving factors.

This just quickly discusses again what we did for the qualitative assessment or the operational expression. We relied on what they provided to the extent possible, but we may have added where we thought we needed to represent the results in the best way.

Here's an example of one of the summaries, and this is from the INL SPAR-H analysis of NFE 1B. So this is the complex scenario.

And basically what is there is mainly what the HRE team provided. We may have added a little bit more to it. But they identified again for the complex scenario that the main drivers would be stress, complexity, and the human machine interface.

And in the SPAR-H analysis, the HMI – well, the main thing they were actually saying is that it was the misleading indicators; that was the problem. And that in SPAR-H is represented by the human-machine interface.

So those are the factors they identified, and those were the things that were weighted the heaviest in the analysis.

I think it's worth reading through that a little bit, particularly at the bottom. It said – well, it says, the situation is highly complex with much more equipment and indication failures, and steam line break and automatic reactor chip would produce elevated stress levels in the crew.

And then they talk about the main factor being the misleading indicators. It says, due to the main steam line break all primary indications of the steam generator tube vector are massed. There is sufficient time for the crews to identify and isolate the ruptured steam generator if they promptly identify it, given the fact that the SGTR is masked, this is unlikely.

So it seems to me at that point they are saying timing is going to be a problem. And I want to bring that up now, because there is a little discrepancy there when you get to the actual comparison.

In terms of negative influences on the execution part, they felt stress was the main factor that would have an influence there.

Positive influences, the SPAR analysis identified available time, experience and training procedures, fitness for duty and work process as nominal.

Now in the SPAR-H analysis nominal essentially means good, but it could be better or it could be worse, okay?

Now I think one thing to note there is

that what they are doing is, they identified available time as nominal. Yet in their discussion they sort of acknowledged that there probably wasn't going to be enough time.

So just a little bit of break down there, and we'll talk a little bit more about how deciding what is going to be the important PSF, which ones will be positive. Which ones will be negative. And how to rate those can be a fairly subtle and complex distinction.

And then the bottom is the positive influences on the execution part, and basically they think everything is going to go fine. The procedures are good. It's a well designed interface, and there shouldn't be any problems.

And then here is the example of a qualitative assessment. I don't think I will go into that in a lot of detail. One thing to note in the second bullet – we could read through it if you'd like – but the main thing to note is, they did acknowledge, they said work process would be nominal is what they finally decided.

Yet they did acknowledge that if they didn't have good work process that could be a major problem.

But we have to acknowledge that none of the teams really had the information to make good judgment about work processes. They couldn't watch the crews in the simulator. They couldn't really interview them and try to get information about how their work processes work, how their team dynamics would work.

So this particular study wasn't set up very well to allow them to do that analysis, because it's a fairly complex analysis. If you want to make inferences abou8t team dynamics and how the leaders are going to work, and if there are any systematic effects across the teams, then you need to be able to observe multiple teams actually performing the actions.

So we weren't able to do that in this study, and that's a problem any time you are doing an HRA and a PRA I think is to have the resources in time to be able to deserve enough crews, or at least to interview enough crews, to make those kind of judgments.

MEMBER STETKAR: John, on the other side is that the analysis without trying to explicitly account for the particular variability across crews within the scenario, the analysis – I don't want to say should –

could have the capability to account for that likely variability, or the degree of variability – the difference in the degree of variability for different scenarios.

In other words a relatively common scenario where you would expect the crews to both perform relatively well and to see less variability across the crews, might have less uncertainty in the human error probability than a more difficult scenario like this where you probably would expect to see higher variability across the crews; or there might be higher variability.

So although it's absolutely true, there was no way to explicitly account for the observed variability, the methodology might be able to account for that somehow.

MR. FORESTER: I wouldn't disagree – it's a good thing to do and it can be done. But I'm just saying that you need to have –

MEMBER STETKAR: But you're saying the methodologies didn't have the capability to do that.

MR. FORESTER: No, it's not the methodologies that didn't have the capability. It's more the study, the resources.

MR. BLEY: If I could comment on that one.

We looked at that in our analysis and said, given how little we know about these crews, they could perform all the way from guaranteed success to guaranteed failure.

So we ended up defining a set of assumptions and saying, if the crews don't meet these assumptions, which we would have determined if we were interviewing them and observing them, here's the result.

Otherwise it would have been way too broad. And in fact it was that broad.

MR. PARRY: I don't think SPAR-H could actually deal with that.

MEMBER STETKAR: That's my only point.

MR. PARRY: I don't think it can. I think you can explore it using sensitivity studies. But you can't – unless you find the algorithm for telling you how you factor that into which performance shaping factor bin you're in.

MEMBER STETKAR: That was my only point is that the methodology – if the methodology doesn't have that hook, thanks, for the term I was looking for, then it's very very difficult for the analyst to try to manipulate that hook, however you do it, whether it's a factor on an estimated mean or a contributor to the uncertainty about that mean.

MR. FORESTER: Okay, I would like to move on to the next presentation.

CHAIR APOSTOLAKIS: Before we do that, Said is leaving soon. So he would like – he will give us his comments from today, and then we will continue.

MEMBER ABDEL-KHALIK: Big picture, I am concerned about the inherent conflict of interest, or the objectivity of the different models owners.

I also would like to see how this work fits within the big picture. And I think the suggestion that we hear a presentation of the overall plan is very helpful.

I do understand that the purpose of this particular part, this particular project, is to collect data to compare against different models.

I also do understand that there are unique capabilities at the Halden simulator. Nevertheless, I do have concerns regarding how representative is the Halden simulator to an actual control room environment.

And therefore I would be very supportive of the idea of having either similar experiments or a subset of those experiments run at a U.S. simulator.

The fact that the crews are from one

utility, same training, same interpretation of the procedures, raises some concerns for me. Because that may not be consistent with what U.S. crews would do.

It's understandable that you want to remove the variability, so you don't want to take 14 crews from 14 different utilities and do this experiment, because that adds another set of variables.

But nevertheless, this is something that needs to be looked at.

It would be a good idea to look at the correlation in performance for the different scenarios amongst the different crews. I know that they are going to do it, and I think it would be a good idea for us to look at that.

There is also a sort of a variability in the response, even of these crews who come presumably from the same utility and the same training, in their degree of adherence to procedures. And one way to avoid that variability is to come up with perhaps some more challenging scenarios, either moving the steam line break to inside containment, and that sort of keeps them busy, or adding something that would even mask the variability in level variation, like tripping one reactor pool and pump, and that would sort of eliminate this sort of knowledge based action, which if I look at the table, it was really all based on steam general level indication.

I'm also concerned about the realism of the simulation, including noise in the control room; realism of operator response; people being able to manually trip the reactor in an event like a major stream line break, to trip it manually just doesn't seem to be realistic.

There are other variables that would be interesting to find out, things like the fact of day shift versus night shift.

But essentially trying to find out how realistic are the data that are being collected to what one would expect in a control room in a U.S. plant.

Those are my comments.

CHAIR APOSTOLAKIS: Thank you very much.

MEMBER ABDEL-KHALIK: Thank you, Mr. Chairman.

CHAIR APOSTOLAKIS: We will see you in March.

MS. LOIS: Thank you.

(Laughter.)

(Off the record comments.)

CHAIR APOSTOLAKIS: So who is next?

MR. FORESTER: I am.

ALL METHODS, GENERAL TRENDS AND COMMONALITIES

MR. FORESTER: I just gave you an example of one of the summaries we did.

Now I'm going to continue on and show you an example of the comparison of that same information from the summary with the actual Halden Crew data.

And again, this is for the INL SPAR-H comparison.

On the left column here we had the negative factors predicted in the SPAR-H submission. And on the right side we have the corresponding PSF that were identified in the data.

So again the SPAR-H analysis identified the misleading indications as being the primary driver of performance here, and in fact that is consistent with what we saw in the actual crew data.

As Ben had mentioned, it was poor; the indications were poor, because the key ones were not there. It could be considered somewhat poor though because there were alternative indications, the steam generator level mismatch, and the flow mismatch.

The next ones down, we have complexity that was identified by SPAR-H as being an important factor. And scenario complexity was in fact, yes, a strongly influencing factor in the actual data.

And then the bottom one was high stress level. The SPAR-H submission thought stress would be a significant factor. Again, this was the complex scenario.

But this is one of those PSFs that we didn't see much of an effect on, although there wasn't much of a report. And a couple of teams might have said they felt some time pressure, but they said they always feel time pressure.

So stress wasn't really significant. There is no data to support that this was a significant factor in the actual data.

Now in terms of the strength of these, the SPAR-H, the multiplier for the misleading implications, the HMI factor was a 50; the complexity was 5; and the stress level was 2. So clearly, the misleading indications was the main driver of performance from the SPAR-H perspective.

The next slide continues with the negative factors here, but the ones on the right are other additional negative factors that were identified in the actual data. These were not identified as negative factors by the SPAR-H analysis.

So for the crews procedural guidance was important. They had problems following procedures to get to E-3. The fact this procedure problem, the procedure guidance was not identified by the SPAR-H analysis. It was treated as nominal.

But they did note in their text that due to the masking of the steam line break and loss of secondary radiation indications the procedures will not assist diagnosis. Yet it was identified as being nominal. So that's sort of an inconsistency.

On the other side we have training, somewhat poor; that was identified to be the case, again, because they weren't trained on this specific scenario. This was not weighted by – or not counted by the SPAR-H team either. It was treated as nominal.

Same for adequacy of time. We – in terms of the crews responsiveness, the adequacy of time was somewhat poor. And the SPAR-H team had noted that time limitations would likely prevent success. But again it was treated as nominal.

So they picked some things and not others to be the drivers of performance.

MEMBER STETKAR: John, these – the PSFs in the right column, I don't remember the whole matrix, but are all of those PSFs included in the SPAR methodology?

In other words the fact that – on the left-hand side it says not weighted by team slash method. Is that mostly that the team had the opportunity to evaluate them; they just didn't evaluate them correctly?

MR. FORESTER: That could be an issue, exactly what you said, they had an opportunity.

MEMBER STETKAR: So if I'm evaluating methods, this isn't necessarily a condemnation of the SPAR-H methodology that it does not consider these things in some way.

CHAIR APOSTOLAKIS: It does consider them.

MR. FORESTER: It does consider them. They just didn't weight them.

CHAIR APOSTOLAKIS: What is adequacy of time?

MR. FORESTER: Is there enough time available to complete the action.

CHAIR APOSTOLAKIS: I was told earlier that the crews were not aware of this 20-minute or 25-minute – so how would they have any perception of whether –

MR. BLEY: They wouldn't have. But the results that you were shown earlier from the simulator showed that half the crews didn't get done in time. I think that is what adequacy of time is. They just didn't get it done.

MR. PARRY: You're right, it's a misnomer. In SPAR-H it really is a misnomer. It's not a PSF; it's a boundary condition.

CHAIR APOSTOLAKIS: Oh, okay.

MR. FORESTER: But they essentially assumed, even though they say verbally they said they didn't think there would be enough time, by saying it's nominal, they are essentially saying, based on our judgment of how long the diagnosis will take, and how long the execution will take, there is enough time. That's what nominal means; that there is enough time.

MR. BLEY: I am trying to remember, it's been awhile since I used SPAR-H. I know in application this is done. I think the method actually says if you don't – for any of the factors they have – if you don't know then you call it nominal. I think that is the guidance for using the method.

MR. FORESTER: I think you may be right.

MR. JULIUS: There's a new update that says, if you don't know assign it as failed. So if you haven't addressed that or gotten the information in sufficient detail to be able to assess it.

Originally it was a nominal.

MR. BLEY: Jeff, do we know if SPAR-H as it was used in this exercise was using that new guidance or that old guidance?

MR. JULIUS: I don't recall.

MR. BLEY: I don't either. I haven't seen that, though; that's good. I'm glad they changed that.

CHAIR APOSTOLAKIS: When HRA models apply, that an assumption that the crew knows – has some perception of how much time is available?

MR. BLEY: No.

CHAIR APOSTOLAKIS: Well, how can time be a factor then?

MR. PARRY: It's not. It's a boundary condition. It is usually associated with a success criterion for the HFE. It's really a PRA imposed success criterion.

MEMBER STETKAR: There can be a PSF for perceived time pressure.

MR. PARRY: For perceived time pressure; right, that's different.

MEMBER STETKAR: If the operators think that they must do something rapidly, or they think that they have a lot of time –

(Simultaneous speakers.)

MR. BLEY: – that would fall under stress, wouldn't it, John?

(Simultaneous speakers.)

MEMBER STETKAR: It is perceived time pressure versus actual.

MR. PARRY: Right, but it's really, did they feel time pressure. And if they did, for whatever reasons, that could impact it.

CHAIR APOSTOLAKIS: Well, it's perceived. They don't do calculations. But there is a time pressure. They do perceive something.

MR. PARRY: But it doesn't go into adequacy of time in SPAR-H.

CHAIR APOSTOLAKIS: No, it's a different one. But models do have that; remember, they do.

MR. DANG: When HRA methods are adequacy of time, it reflects the impact that we expect on the HEP, not necessarily on performance.

It's true for the operators only the perceived time matters. But adequacy of time matters in the sense that we should probably have a higher HEP if there is no opportunity to recover.

CHAIR APOSTOLAKIS: Right, but all that

depends I think on how they perceive the available time.

MR. BLEY: No, I don't think so. It's your point earlier that if in fact the time it takes to do something, regardless if you know how long you have to do it, if that time is on the order of the time available, then are you accounting for that somehow in the analysis. That's what this is about, I believe.

MR. PARRY: That's exactly what it is in SPAR-H.

MR. BLEY: So some people didn't address it and some did.

CHAIR APOSTOLAKIS: But I think in a lot of the models, the amount of stress that they assumed is very much tied to the perception of time.

MR. BLEY: Oh, yes, that's true.

(Simultaneous speakers.)

CHAIR APOSTOLAKIS: Now, another thing, SPAR-H is primarily used in the significance determination process.

MR. BLEY: Is there anybody from staff here? Because I thought they used it more broadly than that.

CHAIR APOSTOLAKIS: That's why it was developed.

MR. PARRY: Yes, but it's also used when people use the SPAR model.

(Simultaneous speakers.)

MS. LOIS: For event evaluation as well.

MR. PARRY: For event evaluation, yes.

MR. JULIUS: Like the activist sequence precursor program.

CHAIR APOSTOLAKIS: Right, those kinds of things.

So there is a claim that SPAR-H is an HRA model like any other?

MR. PARRY: Yes, in that context.

I think also though be careful, because in the notebooks that are used for the oversight process, those aren't SPAR-H values that are in there.

Those are HEP values that come out of a survey of the different PRAs I believe.

CHAIR APOSTOLAKIS: Really?

MR. PARRY: Yes, in the notebooks.

CHAIR APOSTOLAKIS: Which is what is being used?

MR. PARRY: Which is what is being used in phase two. But often in phase three type evaluations -

CHAIR APOSTOLAKIS: They go by SPAR-H.

MR. PARRY: – they go by SPAR-H, yes.

CHAIR APOSTOLAKIS: Nathan, you want to say something?

MR. SIU: Nathan Siu, NRC.

I don't want to go too far off on this tangent, but just the background of SPAR-H, as you might recall. This was put together indeed originally to support the precursor study models.

But basically it was a compilation of THERP, ASEP and the authors' judgment as to what or how the group performance shaping factors, and try to create a simplified approach.

But I think if they were here, they would say that it doesn't go beyond THERP or ASEP, and so it's one way to represent how to use those in that particular application.

Now it is used in the SPAR models, and it is used for a variety of applications beyond the SDP.

CHAIR APOSTOLAKIS: Well, evaluate the accident sequence precursor, and the SDP, what else is there?

MR. PARRY: Well, if NRR staff were using SPAR models to do a sanity check on a licensee's application, they might use SPAR-H in that context.

MR. SIU: Sure, or if you are looking at a generic issue, if you are looking at a variety of things where you want some internal assessment.

MR. BLEY: Of human reliability.

MR. SIU: Yes, as part of a PRA.

CHAIR APOSTOLAKIS: Well, it's a pretty serious use.

MR. SIU: Oh, yes. Oh, yes.

MR. PARRY: As a confirmatory – it would be used as a confirmatory analysis, not as a final judgment, I think.

CHAIR APOSTOLAKIS: All I know is, we have reviewed ATHEANA several times. We have never reviewed SPAR-H from the point of view of what the agency is using, it should be the other way around. We should be reviewing SPAR-H and not ATHEANA.

MR. PARRY: Right.

CHAIR APOSTOLAKIS: So maybe we should do that.

Yes, we will do that.

MR. FORESTER: Okay, I think I will move along here.

This shows the positive factors, the SPAR-H analysis identifies the ones on the left are what they identify as positive in the sense they picked them as being nominal, and the actual data side, the procedures are not identified as being nominal or positive. They were actually negative.

There was some consistency in that we thought the general training and knowledge based support of the transfer to E-3, but obviously the specific training was not good; that would have been negative.

CHAIR APOSTOLAKIS: So again on the left, available time. Is that external to SPAR-H? I really think they use it. Assume that you know somehow how much time is available.

MR. FORESTER: Their analysts have to look at how much time is available, based on TH calculations or whatever. And they have to determine how long it will take them to execute their response.

And then they've got to make a judgment about whether there is enough time left for diagnosis. Because this doesn't work like a TRC where the time for diagnosis gets inferred based on its attracting other numbers out. They have to make that judgment, and then they decide whether there is adequate time or not. And depending on how much extra time there is they would use different factors.

MR. BLEY: That kind of works as a weighting factor.

CHAIR APOSTOLAKIS: I was always wondering how they could make a judgment as to how good the work processes are.

There is no way. You don't know. I mean unless you go and observe.

MR. FORESTER: That's right; you have to observe. Or they have some trainers and operators you can have discussions with; you can learn some of that type of thing.

MR. SIU: Yes, John, just again, at least for some applications, when you are doing a retrospective analysis you do get that information.

CHAIR APOSTOLAKIS: Well, even then –

MR. SIU: Well, to the extent you can collect it, you would – do ask the question.

MR. FORESTER: So here are just some observations on the comparison.

The SPAR-H analysis did identify two fo the more important negative influences: the misleading indications and the complexity. But some factors identified as nominal were actually negative in the crew data – procedures, experience and training, and adequacy of availability of time.

So there is sort of a disconnect there.

CHAIR APOSTOLAKIS: These were negative for all crews?

MR. FORESTER: Yes, there is a lot of commonality here, so these things were pretty strongly negative for all crews.

CHAIR APOSTOLAKIS: To varying degrees?

MR. FORESTER: To varying degrees, right.

Again, I think we were struck by the fact that some of the potential problems that they were going to see were acknowledged, so they reflected a generally good understanding of what was going on in scenarios, but that wasn't really reflected in how they picked the PSFs to rate as negative and positive, and also some of the judgments about the actual weights.

So there are some subtle distinctions there about the decisions that have to be made in SPAR-H; these kind of decisions can be difficult, and that's one particular area we've identified.

And it's not just the case for SPAR-H. There's a lot of methods that could use additional guidance for how to make those selections and those kinds of judgments. So that is one of the lessons learned from the study I think.

On the other hand they did predict a relatively high probability of failure for this event in the complex scenario, and that was consistent with the results.

CHAIR APOSTOLAKIS: So in terms of what we should do, it's not clear to me if I would use your insights, you know, potential cause acknowledged, but they were treated as nominal.

If I see a SPAR-H analysis in three months, you are telling me I should be suspicious, right? But these guys, they are not our guys, have to make a decision. Should they multiply by 10 and then take the square root, or what?

MR. PARRY: You could probably say that of any HRA method.

CHAIR APOSTOLAKIS: I know. That's where I'm going.

MR. FORESTER: I think our point is that no matter who is doing the analysis some additional guidance for how to make those judgments would be very useful. That's something that could be defended.

CHAIR APOSTOLAKIS: And also at the end it says, well, they did predict a relatively high probability. So it's not just the PSFs. I mean it's also how they calibrate themselves, right? If they say this PSF is six, then they have a high probability or a low probability, but also is relevant.

MR. BLEY: And in SPAR-H that waiting is built into the method.

CHAIR APOSTOLAKIS: Is built into the method. See, my quandary here, I don't know where this is taking me.

MS. LOIS: However this is still the pilot. We believe that at the end of the study when we've gone through the whole STGFs we will have better understanding –

CHAIR APOSTOLAKIS: I'm willing to grant you that, yes. But remember – I apologize. As I said earlier today it would be really a good step forward if we saw some change in the models as a result of the insights that you guys are drawing; not just statements, this model and that model has this and that, advantage or disadvantage.

And especially the developers of the models should have an open mind and say, well, gee, I can fix this. Some of it you can't fix. The quality of the work processes, I don't know what you can do about it. It's better not to mention it actually. But other things you might be able to. It's not clear to me what they are.

MR. FORESTER: Well, we're going to talk about some of our conclusions.

MS. LOIS: Gareth has some preliminary insights.

MR. FORESTER: One of the things we're finding is that additional guidance also for qualitative analysis, the sort of understanding what's going on in the scenarios, and good task analysis, cognitive task analysis, a good understand. The HRA methods don't really provide that in some instances.

So again that's a very specific additional type of guidance.

CHAIR APOSTOLAKIS: Another thing I was wondering about is, in one of the great values – at least that's how it has been advertised – of SPAR-H is that it has been structured more than other methods, so it's easier to use; is that correct?

MR. FORESTER: Yes.

CHAIR APOSTOLAKIS: I mean you don't have to go and look for contexts and all that. You are doing a much quicker evaluation. You don't elicit expert judgments and so on.

Susan, are you still on the line?

MS. COOPER: I am. I'm here.

CHAIR APOSTOLAKIS: Is it possible at some point to have a screening using a method that is more structured of the human actions, and then when you identify the important HFEs, to go to a more sophisticated method like ATHEANA perhaps to analyze those in more detail, would that be something that could be used? I don't know, but it seems to me –

MS. COOPER: Was that a question?

CHAIR APOSTOLAKIS: Well, it's a statement. I know you said no last time you were here.

MS. COOPER: I don't remember, and I'm not sure if I understood the question.

CHAIR APOSTOLAKIS: You said no in a very forceful way, Susan. We do remember.

(Laughter.)

MR. PARRY: George, can I add a comment to that?

CHAIR APOSTOLAKIS: Yes.

MR. PARRY: I think if we are going to adhere to the ASME standard, it's the basis for doing the PRA, I think what it would lead to, there is a statement in there that if you want to reach Capability Category 2, which is everybody's sort of goal, then for every significant human failure event you should do a detailed analysis.

But that would be a lot of events in a typical PRA, because the measure of significance is things like Fussell-Veseley greater than –

CHAIR APOSTOLAKIS: No, I wouldn't go with those.

MR. PARRY: Well, that's one of the things on hold right now.

CHAIR APOSTOLAKIS: No, HFEs, I wouldn't do that.

MR. PARRY: Why not?

CHAIR APOSTOLAKIS: Because they are very conservative.

MR. PARRY: Not necessarily. My problem has been with screening that I've seen too many studies where people have used conservative screening values. They put in point one because somebody says that's conservative, and it's okay to use in screening. And they multiply it together six times and get 10 to the minus sixth for an integrated human error, and yet it's screened out, because at each point it's screened out because they have not done an integrated –

CHAIR APOSTOLAKIS: Obviously if you want –

MR. PARRY: And each one individually comes up with a low Fussell-Veseley importance, or however you want to put the importance, if you treat it as a valve.

CHAIR APOSTOLAKIS: But if you work to

develop an approach that included screening and so on, you would sit down and think about it a little bit. You don't just say, let's screen tomorrow; what can we do? No, those values and that screening was developed really for other purposes, for 5069 and so on.

Here you have a different thing. You will think about. You will think about the issue that John just raised, and you will try to do it.

I'm trying to get away from this problem. ATHEANA is accused of being too detailed and taking forever. On the other hand it has some excellent features.

This one seems to be much faster and so on. Not all HFEs deserve a detailed treatment perhaps. There must be a way out of this. We can't just stay like that forever.

MR. BLEY: One thing that would help a lot is if people doing HRAs by any of these methods developed a good qualitative description of the things they are trying to examine.

CHAIR APOSTOLAKIS: Absolutely, yes.

MR. BLEY: And if they do that, a lot of the kind of things, like maybe we were pointing ou8t there that weren't thought abou8t very well, kind of get forced into being thought about. And that's not a function of – not that that's a function of the process for doing HRA, and a lot of people use methods without that. They take an HFE, take a couple of performance shaking answer, and get an answer.

CHAIR APOSTOLAKIS: Gareth just told us that in the SDP, we knew that of course, but he refreshed our memory, there is a phase two where you use this, and then we go to phase three if necessary.

Why couldn't that be applied to human error? Do quick calculations, and then if necessary, whatever necessary is, we have to define it as a community, then I move on to something more detailed, and I have an interaction, why couldn't we –

MS. LOIS: Our concern is, we are not quite sure that the SPAR-H use results are consistent. And there is an indication that may be underestimated.

So if you use this as a screening analysis you have to develop a guidance in how you really use SPAR-H as a screening tool.

CHAIR APOSTOLAKIS: Maybe I misspoke. I'm pushing the idea of screening and then detailed analysis. I'm not pushing –

MR. PARRY: And I don't think anybody disagrees with that.

CHAIR APOSTOLAKIS: Okay, all right.

Now the other thing is, has anybody ever tried in a PRA to use two different models for the same HFEs?

MR. PARRY: For the same HFEs?

(Simultaneous speakers.)

CHAIR APOSTOLAKIS: – failure event. I have somebody who does it with ATHEANA, somebody who does it with SLIM or whatever.

MR. BLEY: We are trying it here.

CHAIR APOSTOLAKIS: Yes, but this is a research project. I'm talking about the PRA.

MR. JULIUS: I've done it, and I've seen it done in SEP situations. A lot of times where a method won't fit, we'll show that, what if you tried SPAR-H, what if you tried the EPRI cost-based decision tree, what if you used the EPRI HCR to see what the different methods –

CHAIR APOSTOLAKIS: And what did you find?

MR. JULIUS: – to see if maybe there's some sort of loci of HEPs or you know –

CHAIR APOSTOLAKIS: And what did you find, Jeff?

MR. JULIUS: Typically the ones that you're looking at are at the bounds and outside of the normal modeling range, you know, the methods anyway. So there were – some places there were wide variations. At other times they were actually relatively close but then you were close to a high HEP anyway. You know, you were somewhere like between point one and point three.

MR. PARRY: Yes, I'm not sure that these are – I think you are right, Jeff, usually the things that have been done in that area are where none of the methods are strictly applicable.

MR. JULIUS: That's right.

MR. PARRY: So it's not really a good comparison.

MR. JULIUS: It's not a real test.

MR. PARRY: It's not a real test.

CHAIR APOSTOLAKIS: Because that would be another way of handling the insights that Vinh showed us. In other words, if I do it – I know I can do, but let's say three different ways. And I get three different results with some spread and all that, maybe based on judgment. Then I can come up with something composite that will reflect presumably some of the concerns that you identified.

But that I think is a dream. I don't think anybody would ever spend the money to do this.

But the screening followed by more

detailed evaluation would appear to be something –

MR. BLEY: The only thing close to that I know of is that people have done a PRA and then some years later have gone back and changed their methods and reevaluated the same events.

CHAIR APOSTOLAKIS: Well, that would be useful to see.

MR. BLEY: But mostly with licensees who haven't published that or shared it.

MR. PARRY: And also I think they may do it as part of a broader revision of the PRA, so some of that detail gets lost I think.

Do you know of any cases like that, Jeff? Because I know you have been involved in converting some PRAs to the EPRI calculator.

MR. JULIUS: You're right, Gareth, like Diablo Canyon was the first one we did, and they converted their SLIM over, and – but it was done as part of an overall update where they did data changes and success criteria changes, and we didn't focus on the individual, the overall impact.

I guess what we did see, because they didn't want to run before – with the old HEPs and the new HEPS, and we thought it was some shuffling around. But I think generally the ones that were high before were still high, and the ones that were low – you know the relatively ranking was –

MR. PARRY: Which is encouraging in itself.

MR. JULIUS: Yes, that's just my recollection. We didn't do a detailed comparison.

MS. COOPER: I wanted to make a comment to try to respond to George's comment earlier.

CHAIR APOSTOLAKIS: Yes, sure.

MS. COOPER: I think what we are trying to do now in the joint EPRI-NRC fire HRA development project may actually end up being close to what you suggested, and that is that at least part of that, not all of that, but part of that is involved trying to bring some of the ATHEANA concepts into sort of a screening type approach.

I mean the overall idea that we are heading for without saying too many things in advance of finishing any work, but the overall idea is to try to relax some of the conservatisms that are in the current guidance provided, and make it a little bit less conservative with what we're calling sort of a scoping HRA, some of which is based on sort of the EPRI methods, and some of which is coming from some ATHEANA ideas.

But then the ultimate sort of direction is if none of that is sufficient, for whatever reason, you still don't like – the result is still higher than you'd like and you'd like to see if you can do some more analysis and reduce the HEP, then you go to one of the detailed methods.

CHAIR APOSTOLAKIS: That's good. That's good.

MS. COOPER: That's a work in progress right now.

CHAIR APOSTOLAKIS: I guess what I'm trying to do is remind you what the outcome or goal is here. Ultimately that's where we – we want to go someplace where this large number of models and all that, we know what we want to do. Just as an Agency. It really looks bad, Gareth. The situation we have now is very bad.

MR. PARRY: Well, I doubt if it's very bad. I think it's workable. But I just thought of something that goes counter maybe to this screening and detailed model. And that is, that might be okay for a base PRA. Sometimes once you start using it, then if you are using it to explore a certain issue, that screening value that you had before might now be a very important value. And I think if you are using a different method to analyze that, you are going to have – the whole model is going to be somewhat inconsistent I think.

So you've got to be really careful. Seeing this with five PRAs, for example, with a base five PRA – and I'm not talking about HRA now – you could develop a model which gave you a satisfactory answer, maybe, even using conservative inputs.

But now you are going to look at it for doing tech spec evaluations. And then you find that, oh, this conservative evaluation I made is really screwing up my analysis.

CHAIR APOSTOLAKIS: Then you revisit it, yes.

MR. PARRY: So you've got to be a little careful about how we set this stuff up.

CHAIR APOSTOLAKIS: PRAs using – maybe screening is not the right words. Maybe a phase one, phase two, phase three, or first cut – you know, we do that all the time. I just remember in the old days, PRG, they did two or three very detailed PRAs, and then the next plant, well, it's more or less similar to that plant. Let's do a point calculation first.

MEMBER STETKAR: Well, George, be careful, because you were at PLG before we learned that it wasn't a good thing to do that also. In the late `80s we kind of discovered that that probably wasn't what one wanted to do.

MR. PARRY: Like when Zion ended up --

MEMBER STETKAR: Yes, we finally figured out that Zion didn't look like South Texas for example.

(Off the record comments.)

MEMBER STETKAR: But there was some move 25 years ago to do that sort of thing.

CHAIR APOSTOLAKIS: But don't tell me now that the idea of a phased approach is entirely new to PRA. I mean come on, we do that all the time. You do it a certain way –

(Simultaneous speakers.)

MEMBER STETKAR: the only problem that I've seen in HRA in particular, because of the difficulty – take an analogy. In common cause failure modeling for example a lot of people will simplify the process and use a simple beta factor model, which is numerically conservative for a highly redundant system; but it's a well defined model, and as long as you define the population of equipment you have constraints on that, and you can measure how relatively conservative you are.

In HRA there have been many many many

examples where people have followed this process of doing a so-called conservative screening analysis and saying, well, none of these human actions are important, but have lost the real context of the human reliability analysis because of the way that they have defined the human failure events in the model, and then used the methods to so-call assign a conservative number to that particular little box.

So when you go back and you look at – you ask somebody to do a simple example of, requantify your models setting all the human error probabilities at point nine, and show me the cut sets. You have cut sets with seven or eight different products of human errors, within – that must be accomplished within 30 minutes – and they are screened out because the composite conservative screening value is 10^-6, simply because of the way the model has been constructed.

It's a different way of screening. So when you are talking about screening, or this phased approach, you need not just to look at the – whether it's SPAR-H or ATHEANA or – I don't care about the calculator – it's how the human reliability is integrated into the risk assessment that is equally important.

CHAIR APOSTOLAKIS: Maybe we don't mean the same thing. By screening I don't mean that you find out. I mean it does not deserve –

MEMBER STETKAR: Detailed analysis.

CHAIR APOSTOLAKIS: – more detailed evaluation.

MEMBER STETKAR: My whole point. People have not done more detailed analysis on any of these actions, because they have determined that they are insignificant, because it's point one to the sixth. So therefore none of these actions require any more detailed analysis.

CHAIR APOSTOLAKIS: Well, they did it wrongly, but the idea is not bad.

MEMBER STETKAR: The concept is not bad; it's only how it's implemented.

MR. SIU: It will take work to develop a basis for assuring yourself that your phase one analysis is indeed appropriate. So for example collecting data, comparing analysis results against that, is something you would have to do.

CHAIR APOSTOLAKIS: It's important I think when you say things like that to bear in mind what we do today.

Today we have groups of people who develop models who are pushing them, and there is absolutely no cooperation. And I want to get away from that. If one method can help us do something weakly, maybe it has some value and can use it there. And then I use something else which is much more sophisticated, which means it takes much more time to do in resources, to focus on something that is more important in some sense.

Now how I implement that, I didn't say you can do it in a day. But at least if we are open to that, and we are learning from the insights that you are gaining now, maybe we can start moving that way.

We'll do it right. Of course we have to do it right.

MEMBER STETKAR: That's one of the reasons why I asked early this morning, was there any feedback from the HRA teams on the level of effort required for each of their analyses, because if the application of SPAR-H, if SPAR-H captures the most important performance shaping factors as evidenced by the experiments, if the hooks are in there, using Gareth's term, if the hooks are in the methodology, and it can be applied with relative efficiency, then it satisfies a lot of the criteria you are looking for.

How the particular HRA team, their

capability to apply the methodology correctly is a different issue. But identifying efficient relatively robust methods I think is something that would support what you are looking for. As opposed to the other end of the scale, the best possible but perhaps very very resource intensive method.

CHAIR APOSTOLAKIS: Exactly what I'm saying.

The other thing, though, let's not forget that just capturing the PSFs is a very important first step. But then when you translate things to probabilities is also very important. I think you are not investigating that now?

MS. LOIS: Yes.

MR. PARRY: Yet. Although we've seen enough evidence to know that there could be better guidance in that area.

MS. LOIS: To get to the point of the level of effort needed, we did not ask that question, but from the analysis provided, it's apparent that those teams that spent more time in thinking for this scenario provided a better analysis.

And it seems that one of the take-home results is what we were talking about, developing guidance as to how the method should be applied. And the level of effort that you spend on human reliability may not be a good excuse for the results that you provide. I mean that commensurate to the application.

MEMBER STETKAR: There may be another variable, though, there. If all of the teams were doing this part time, under their own funding, perhaps some teams had some vested interest to spend a lot more effort and time, and had more internal resources to just prove that their methodology was the better methodology.

MS. LOIS: Regardless of the reason for which you didn't spend the resources, I mean you may – I am poor and don't have resources, but the quality of the analysis shows that resources are needed.

And the idea that you can do HRA on a back-of-the-envelope calculation, and you can just go and use PRAs, and you know, use the multi – various factors to multiply up and down, it becomes really obvious that that is not the way we should do HMA.

MEMBER STETKAR: Let me flip that and put it in kind of a commercial sense. And that is, given a budget of $2,000 and four days to produce the best HRA using my methodology that I can produce, for $2,000 in four days; and that's it; that's all. Then give me the results. That is a measure of efficiency. It's a different problem.

You then provide an incentive, and if you don't produce the method – if you don't produce the results in four days you fail. Your methodology obviously cannot support this.

MS. LOIS: So we believe that after this study we will be able to create tools that will allow people to use their method effectively and efficiently without tremendous expenditures of resources. We believe.

CHAIR APOSTOLAKIS: So you don't think that it's possible to come up with a combination of methods?

MS. LOIS: Oh, no, we believe that too.

CHAIR APOSTOLAKIS: But let's not forget, this is a typical discussion it seems to me among researchers. Let's not have – what do we call it? – the perfect is the enemy of the good enough.

MR. PARRY: Something like that. That's a good enough statement.

(Laughter.)

CHAIR APOSTOLAKIS: You should always bear in mind what the situation is now. So can we make a step or two forward, maybe not reaching the perfect state, but at least improving on what we have now, which I think is not – I mean if it was good enough we wouldn't be doing this.

MS. LOIS: We believe in collaborating with EPRI I think we are getting there as an Agency. And this –

CHAIR APOSTOLAKIS: The other thing I am trying to make with my comments is to sensitize you to the fact that ultimately you want to respond to the SRA, which says, a model or a suite of models.

So all these insights, all these are great. But you should always ask the question, so how does that help me to get there?

You want to talk now?

MR. PARRY: I could do. I mean most of what I was going to say has been said, so I can talk really quickly if that's okay with you.

CHAIR APOSTOLAKIS: Okay, then after Gareth, you have more presentations?

MR. FORESTER: If you want them.

CHAIR APOSTOLAKIS: If I want them?

MR. FORESTER: We had a couple of things. You know Jeff Julius was going to talk about the ATHEANA comparison, just another example, and we could do that. And Vinh –

MR. JULIUS: But in summary, I was just going to say that it's too detailed and takes forever.

MR. BLEY: Thanks, Jeff.

MR. FORESTER: And Vinh was going to talk a little – give you a little bit bigger picture of the comparisons across the methods.

And then we had one presentation –

CHAIR APOSTOLAKIS: Is it possible to finish by 5:00 o'clock?

MR. FORESTER: Sure.

CHAIR APOSTOLAKIS: It's always possible.

MR. FORESTER: That's up to you.

CHAIR APOSTOLAKIS: But I don't want to have somebody feel that they have something important to say and we didn't give them time.

MR. FORESTER: I don't feel that way. I'm sure Jeff would be willing to skip the ATHEANA part.

MR. DANG: Which is in the report.

MR. FORESTER: Which is in the report. And what Vinh has is not in the report.

MR. DANG: And it's short.

MEMBER STETKAR: Yes, that looks interesting.

MR. FORESTER: And then my talk was going to be a little bit about HEPs. It can wait until the next round.

CHAIR APOSTOLAKIS: Yes, it's a little too soon to talk about the HEPs, isn't it?

INSIGHTS FROM THE PILOT STUDY

MR. PARRY: All I really wanted to do is to summarize –

MR. FORESTER: Let me get it up.

CHAIR APOSTOLAKIS: Do you have slides?

MR. PARRY: Yes, we have slides.

Actually Gareth is not up here yet. Let's see.

(Off the record comments.)

MR. PARRY: Okay, good.

What I wanted to do is to briefly – I was going to recap the summary of the approach, the comparison. But I think you understand that well enough.

What I wanted to do is just remind ourselves of some characteristics of HRA methods. And I am specifically talking now about the quantification part of those methods that affect the comparisons that we were trying to do; then summarize the insights; make some observations on the pilot; and then draw some conclusions.

So if I go forward to slide #5, it's

called characteristics of HRA methods, let's remind ourselves that basically most of the HRA methods that we are dealing with here have been developed primarily to support PRAs, and one important thing about PRAs is that the accident sequences in which we embed the human failure events represent effectively a discretization of all the possible scenarios, which means that there is a lower level of variability that we don't address, and it's captured – somehow it's all embedded in the definition of the sequence, and the HFE.

And what I wanted to point out here specifically is the crew-to-crew variability. Because that's really the thing that we've been able to see here.

Because mostly the way the plant scenario was developed was controlled. But in this context crew-to-crew variability is basically an aleatory factor, which means that it gets factored in, and you put in a probability, which is the probability of – addresses things like the probability that the particular crew that is on duty is a random variable. So the only thing we feed in there is a single probability. It's a single HEP.

So in that context we don't have a way of addressing crew-to-crew variability as such, except in some specific methods.

And some methods, like the TRC method that is based on simulator trials for example, explicitly does portray the crew-to-crew variability, because that is what drives the shape of the curve.

MEMBER STETKAR: But it's only a single curve. There is not uncertainty about that curve to really show a range of variability.

MR. PARRY: No, and I'm not even addressing all the problems of HR –

MEMBER STETKAR: It is a mean curve.

MR. PARRY: It is a curve that represents crew-to-crew variability, and you use it when you can identify a time window, and where that curve crosses the time window is the probability of the human error probability.

Other than that there's like ATHEANA and MERMOS and I heard from John this morning, maybe SLIM, has the capability of addressing crew-to-crew variability in the way that they calculate the mean human error probability that goes into the models.

Some of the methods, though, like SPAR-H, I think the CBDT, other methods like that, do not address crew-to-crew variability.

And the way you address it, I think, is you look for some sort of average characteristic of the crews. So you look at average training; average experience.

So it's very difficult to handle this crew-to-crew variability in those particular methods. Now what you can do with that, though, you could – if you were using the PRA as an exploratory tool to find out how you could improve your plant for example, you could look at the sensitivity to crew-to-crew variability and try and figure out whether there was something you could do to try and reduce that.

So I think even those methods have some facility for exploring this concept. But I don't think – they are not set up to deal with it.

CHAIR APOSTOLAKIS: But even the PRCs that are developed in the EPRI approach, they are applicable only up to a certain –

MR. PARRY: Yes.

CHAIR APOSTOLAKIS: – which I think you can translate to the words that Dennis used earlier. If they are comparable times, if the time available is comparable, it's the same idea. Because if they tell you the available time becomes much larger than this, then you have to switch to another method.

MR. PARRY: And that was the philosophy behind the CBDT, for example.

CHAIR APOSTOLAKIS: Right. So already the idea is there of focusing on time, or certain class of problems, and then moving to something else, is there.

MR. PARRY: Yes.

So that's why – I guess the reason I brought that up is that that is to some extent a complication when by comparing with the results of the simulator studies. Because in fact one of the major insights we got from the simulator studies is that crew-to-crew variability is a significant factor.

CHAIR APOSTOLAKIS: But you didn't run any exercises where the available time was much longer?

MR. PARRY: Probably some of the events later on in the success path are like that.

CHAIR APOSTOLAKIS: Are you going to focus on those later?

MR. BLEY: We have analyzed those.

CHAIR APOSTOLAKIS: Sorry?

MR. BLEY: We have analyzed those already.

MR. PARRY: The HRA teams have analyzed those. We haven't assessed them with respect to what we have learned from the simulator trials.

MR. FORESTER: We are going to do that,

though. We are in the process.

CHAIR APOSTOLAKIS: That would be a useful thing to see.

MR. PARRY: Okay, so that's one point I wanted to make.

Another point I wanted to make, which I think was already made, is that all the methods, just looking at them overall, they did identify some of the important driving factors, but probably none –

(Simultaneous speakers.)

MR. PARRY: It would be pretty sad if they didn't.

But I think what we also saw is the fact that – and I don't mean PSFs – the factors in the general sense that drive the variability between the HRA team predictions.

I think Erasmia mentioned, really, a lot of it was to do with the depth of the analysis to develop the qualitative understanding.

Some of it I think was also driven by assumptions based on the fact that they didn't know much about what the crews were really like. Certainly in a couple of cases it was for the so-called simple case of the simple steam generator tube rupture for example.

A lot of it was based on what people assumed about when the reactor was going to be tripped. I mean some teams said, okay, the standard that we saw – the printout that we saw from the computer said they were tripped at six minutes and 53 seconds. That's the time we will take at the trip, which doesn't allow for variability in that factor for example.

Other things that affected the variability was in fact the PS actually used by the methods. In some cases I think it's the definition fo the PSFs that we use by the methods. There was only one method that was used twice, and that's SPAR-H. And I think looking at the results, the way that the observations, or the way that the thoughts about what was complicating the scenario was fed into the PSFs was different, which tells me that in fact the PSFs are not defined very clearly in terms of what should be captured in those PSFs.

So what that leads us to believe is that – if we can move on to the next slide – is that all these methods could do with an improvement in the guidance on how to use them, particularly in how to develop the quantitative analysis, and I think specifically in how to address the strengths of the

PSFs, and in the case of SPAR-H, maybe even how you factor things into the right PSF.

MEMBER STETKAR: When you say, strength of PSFs, you mean –

MR. PARRY: The impact on the HEP.

MEMBER STETKAR: – how relatively important –

MR. PARRY: Relatively important; the weights if you like.

MEMBER STETKAR: The weights, okay.

MR. DANG: And the rating.

MR. PARRY: To me that's the same thing I think.

CHAIR APOSTOLAKIS: Well, rating is different usually. That's important too.

(Simultaneous speakers.)

MR. PARRY: Okay, I guess I was thinking of the SPAR-H model where the weight is in fact –

MR. DANG: The weight is always the same.

MEMBER STETKAR: But there are differences. For example, when I talk to people, tying my shoe is a good example. If I have a PSF procedure, and a PSF for experience, procedures are terrible. I don't have a written procedure on how to tie my shoe, but that is not very important. So although it could be – you still do? That's sad.

That's in some sense what I was trying to get to an understanding of what you meant by strength of the PSF, because in my sense, the fact that I don't have a procedure isn't very important –

MR. PARRY: Because it's compensated by training?

MEMBER STETKAR: Yes, because this is not a proceduralized type of activity that I am talking about.

MR. PARRY: And I think, too, it's the interplay between the different PSFs as well.

So I think there are – there is room for improvement on all the methods that we believe we saw.

For example I reviewed the CBDT, and I would have analyzed it differently from the way they analyzed it. And actually that's a bias maybe, but it does show that these methods ought to be I think reproducible to some extent, and translatable from one analyst to another, if we are going to use them in a regulatory context. I think we have to have some sort of –

CHAIR APOSTOLAKIS: And you know this infamous ISPR exercise. We have to put it to rest at some point.

MR. PARRY: Yes, we should, definitely.

But I think there are more problems with ISPR –

CHAIR APOSTOLAKIS: They have huge problems.

MR. PARRY: They have huge problems.

CHAIR APOSTOLAKIS: But it was disturbing.

MR. PARRY: Yes.

So that's one of the things that we need to do.

Just some observations on the pilot. I think – and this really is perhaps one of the reasons why we are not focusing on the HEP specifically, because both of these scenarios, I think, are untypical of the ones they are modeling in PRAs.

And as a result of that, I think that we did see quite a large range of the HEPs that were reported to us.

A lot of that was to do with, I think, the fact that really understanding how the crews would operate, even though we understood they were trained and what their procedures were, I think it led to some assumptions that were made that made big differences to some of the answers.

Let me just make a couple of comments on the two HFEs. The HFE 1A as we talked earlier involved the crew in making a decision on when to trip the reactor. And that was the big factor. Because if you left it just a little too late, it meant you didn't have enough time to meet the success criterion, as was identified in this particular HFE.

In terms of HFE 1B, the complicated steam generator tube rupture, obviously this is a really low probability scenario, particularly in that it involved a steam line break followed by a tube rupture and oh by the way you have an instrumentation failure.

So what it requires, I think, was an appreciation for how the crew would deal with incomplete information, if not possibly conflicting information, conflicting would be, well we didn't see any radiation, but we do wonder why the steam generator level is going up for example.

As it happened, of course, I think most of the crews actually realized they had a steam generator tube rupture because of all these other indications. But in the way that the procedures were set up, it didn't lead them naturally to using those indications.

So I think that our conclusions as far as the pilot goes is that we think it has demonstrated the value of performing the comparisons of predictions with empirical data, particularly on the qualitative side of things.

Really quantitatively the jury is out on that one, but I don't think it's very useful. Because if we want good statistical data from these experiments we've almost got to force the situation so there are failures.

And that's not really very helpful I don't think. So we do – we are able to I think make good value of this, and we understand – I think we've learned quite a lot from doing the pilot in a way that we can define future exercises much more precisely so that we can in fact control the information that we get a lot better.

And that is really all I wanted to say. And I'll leave it open –

MS. LOIS: And we have documents needed for the loss of feedwater, we haven't – the analysts did not analyzed loss of feed, and therefore we have opportunity –

CHAIR APOSTOLAKIS: When is this exercise going to finish?

MS. LOIS: We believe that we are going – the results of the whole analysis including loss of feed in 2009.

MR. PARRY: But we didn't have – again, the loss of feed scenarios were already run. So we don't have the chance to go and redefine those scenarios. They are what they are.

MEMBER STETKAR: This is a total loss of all feedwater.

MR. PARRY: I believe it is.

MR. BLEY: I haven't seen it.

MEMBER STETKAR: Oh.

(Simultaneous speakers.)

MS. LOIS: We are not ready to discuss this.

CHAIR APOSTOLAKIS: So Vinh is next?

MR. DANG: Can be.

STUDY METHODOLOGY, ORGANIZATION, PARTICIPANTS

MR. DANG: This is a short presentation. Let me find it first, and then we'll start.

Basically it deals with the method to method performance. We've been talking about how the method did versus the data. And we expected that we wouldn't be able to get out of this room without answering the question of how did the methods do relative to each other.

As a reminder, the three parts are: how do they do on the predicted level of difficulty (that's

AGP); that's going to be addressed in a follow up. I have a few more comments about that coming up about why that is so.

The second is, identifying correctly the factors that lead to this difficulty.

And the third is, identifying correctly the features or specific characteristics that underlie these factors.

These are important, because we think that the HFE only has a value if it's correctly – if the contributors are correctly identified.

In addition if you are going to use the HRA result in the sense of what is the weakness there, if you are pointing to the wrong factors, you are going to get the wrong things fixed. So that is quite important.

One of the reasons that the predicted level of difficulty will be addressed in the follow up is that it really has to be done for a set of HFEs so that method performance is being assessed on a more – well, maybe not representative, but a broader set of HFEs that perhaps represent a more diverse set of unreliability factors, and not just the ones that you have in 1A and 1B.

In addition, because of the sparse data, we think that by looking at more HFEs as a group we can do some rank order comparisons about whether or not they are getting the relative difficulties correct, even if the empirical data only suggest tendencies towards failures, and not actual failures.

So the method performance and driving factor identification is what has been the focus in the work so far.

Does the method identify the contributors to HFE difficulty, in terms of the high level areas, meaning in terms of topical areas like procedures versus training versus indications?

Although if we only look at driving factor identification, we need to interpret these results with caution because they can be misleading if you don't look at what specific features and characteristics have been identified.

So you can say procedures are poor for the wrong reason. Again you got the general attention right, meaning that people were looking at procedures. But they will be fixing the wrong part of the procedures.

So this last issue is this method performance on the features operational issues that underlie the driving factors. And we have done this for each method.

We looked at the issues, identified them for the driving PSF and did the matching. We looked at the operational expressions. We've shown a number of these things in the – well, in John's presentation. In the distributed slides on the ATHEANA comparison, and in the report.

The method-to-method comparison at this lower level has really not been done at this stage – has not been done at this stage, sorry. The issues that are being identified depending on the method are being identified at different levels, also, meaning at different levels of detail.

So method-to-method comparison about these insights is a little bit difficult. How much value is the insight? How correct is it? And so on.

Finally it should be said that in normal practice, and maybe usual PRF applications, this question is not really asked of the HRA analyst. They are not used to giving this answer.

So we want to make sure that they understand, and I think the pilot has contributed to this, they understand what we are asking them when we say, give us operational issues.

I think by telling us – by telling them in

1A and 1B what we considered to be operational issues, then when they submit future submittals, they can explain what they think are the operational issues in ways that are more congruent with the ways we have been doing, or the way we have been thinking about it.

The general trend is that the more recent methods place much more weight on these underlying features and operational issues, and the positive thing is that it makes it easier to compare against the simulator data because these types of predictions if you will are more concrete.

Or if you say that step is poor, it's very easy to tell whether there were problems with performance for that step as opposed to the general procedure.

Okay, this is my last slide. It represents in tabular form the results of the complex case.

The left number before the slash is the negative factors, and the right are the positive factors. I know I'm putting up a lot of information, but I'll try to walk you through it.

So if we take the first one, ASEP NRC, they made a total of six predictions of driving factors, six negative factors. Five of these were identified in the data, so that was pretty good. One was predicted but not observed.

And one of our observations is not reflected at all in their predictions. Now it may be that they missed it, or th8at their method doesn't address that issue.

On the positive factor it's similar.

Now the caution when one wants to compare these things is, if you look at this, you say, well, ASEP NRC, they make six predictions; they got five right. ATHEANA presumably with more effort made five predictions, got four right. It looks about the same.

But if you look down at the underlying predictions, you see that first the decision component in the ASEP analysis misses the negative driving factors on the decision. So they put the driving factors, the negative driving factors, on the execution component, which did not turn out to be the main issue in this complex case.

CHAIR APOSTOLAKIS: But isn't there an issue here? Is it the methods that resulted in these, or the people who implemented them?

MR. DANG: Both.

MEMBER STETKAR: It's really interesting that you brought that up, George. Because I've used

ASEP quite a bit in the past. And I have a little routine on my computer that does interpolations and all that type of thing.

I noticed that the NRC ASEP team in particular for the complex case abandoned the time reliability correlation. They didn't use it. They focused for some reason instead on just the implementation phase, and I don't know what they did.

Had they used the time reliability correlation in ASEP, and used the high curve which you could argue is bad procedures, difficult training, and so forth, they would have come up with something like a point five for failure to correctly diagnose the situation within about a 12 or 13 minute time window, which is roughly, from what I could discern, which is pretty doggone close.

That says that perhaps it is the HRA team, not consistently applying the method, rather than a shortcoming in the method. And that's the only one that I tried when I was going through this whole thing. And I thought, gee, that's very very curious.

MR. DANG: And we do recognize this to be a methodology issue. Basically the only way to deal with it is to control for the analysts, meaning to have several teams using the same method. But that's a different –

CHAIR APOSTOLAKIS: Who implemented SPAR-H?

MR. DANG: Excuse me?

CHAIR APOSTOLAKIS: Who implemented SPAR-H?

MR. DANG: I know it was one –

CHAIR APOSTOLAKIS: Not the whole lab, I mean who?

MR. DANG: The persons?

MS. LOIS: SPAR-H was – we don't – we haven't presented any results – it was also an NRC team that participated.

So we say SPAR-H INL. We have another SPAR-H analysis by the NRC. And the two analyses are very different in terms of insights as well as numerical results.

MEMBER STETKAR: That's the only method that was tried with two different analysis teams.

CHAIR APOSTOLAKIS: Can you tell me how well MERMOS performed? I still don't understand the 3-1 3-0. Which numbers are the important ones?

MR. DANG: This number is rather important because it's the number of predictions that were correct. So three out of three is good, but they did not identify many driving factors. It's a feature of their method.

CHAIR APOSTOLAKIS: Aren't they the ones who claim that their method is completely different from everybody else's?

MR. DANG: It is, and as a matter of fact in the comparison it's very difficult, because of the structure of their method, to use the same driving factors. And in fact our comparison for their method has focused on the operational issues.

And on the operational issues they do quite well. I do have a transparency on that if you would like to see it?

CHAIR APOSTOLAKIS: No. You mean a slide?

MR. DANG: Yes.

CHAIR APOSTOLAKIS: Oh, okay, a slide is okay. I thought you wanted an overhead projector.

MR. DANG: Okay.

(Off the record comments.)

MR. DANG: So this is very close to what the memo produces to describe the dominant way in which the HFE is expected to fail. And one descriptor is that the system, and by which they mean the crew and supported by their procedures –

CHAIR APOSTOLAKIS: And see, that's what confuses the hell out of me. Everybody seems to meet – to be using the procedures. So what is it, just a renaming of things? Because they make a big deal out of it, that it's not the crew; that it's the system which means the crew, the procedures, and whatever else.

MR. BLEY: That includes some automated procedures normally. I don't know if they automated these.

CHAIR APOSTOLAKIS: But it seems to me that every single method uses all that stuff.

MR. DANG: I think it is beyond the scope of this last half hour to try to answer your question.

I do think that their way of approaching the problem is quite different. It really is not the same.

The language is very opaque, but it does reflect a different way of thinking about the problem.

CHAIR APOSTOLAKIS: Okay, so when are we going to learn that?

MR. DANG: You can ask EDF.

MR. BLEY: You can buy it.

MR. DANG: We are talking about free licensing.

MR. BLEY: Are they really?

MR. DANG: Yes.

MR. BLEY: Boy, that's a big change.

MR. DANG: And English language courses.

MR. BLEY: Really?

MR. DANG: Yes, I don't know if that will be free.

CHAIR APOSTOLAKIS: Is the APR PRA using MERMOS?

MR. DANG: Yes, all the EDF PSAs are using MERMOS.

MS. LOIS: Not for the U.S. plants.

CHAIR APOSTOLAKIS: Not for the U.S. plants.

MR. DANG: They are PSS.

MS. LOIS: But my understanding is that the EPR analysis, PRA that is going to be part of the application for the – to be reviewed by the NRA is good in SPAR-H.

MR. DANG: Oh, I wasn't talking about the EDF PSAs. So the 900, the 1,300, the N–4, all of them are using MERMOS for full complete applications submitted to the regulatory authority.

CHAIR APOSTOLAKIS: Okay, I guess we can learn about it.

All right, so what is your conclusion?

MR. DANG: My conclusion is, you will see more of the method to method comparison. I wanted to give a flavor for the difficulties of grading the method on this kind of complex performance of the methods, given that they address things at different levels; that they involved different levels of effort; that they have different focus.

We are going to try to roll it up. I guess one of the things that we find very positive is that the HRA teams themselves are finding this type of assessment, and the opportunity to compare against simulated data, to be very valuable for their future consideration.

So they appreciate this chance to see how well they do against something real.

Yes?

MEMBER STETKAR: In this sense of – something I just thought about – is would you characterize the HRA teams as simply practitioners using these methods? Or is it closer to characterize them as champions of the method?

For example, Dennis and Susan apparently participated in the ATHEANA and they want to show the ATHEANA is the best methodology in the world, as would presume the MERMOS people.

What about the SPAR-H?

CHAIR APOSTOLAKIS: Blackman was the

developer.

MEMBER STETKAR: That might be a way of kind of normalizing – people say, well, it doesn't look like this team put a lot of effort into their analysis. But if they are the champions of the method –

CHAIR APOSTOLAKIS: The champions are long gone.

MEMBER STETKAR: That's true.

MS. LOIS: However, the NRC's SPAR-H team on purpose they did the analysis the way they exercise it in the field.

MEMBER STETKAR: They are more practitioners.

MS. LOIS: They are practitioners, and they say, we just want to demonstrate how we would do the analysis if it was part of our task.

CHAIR APOSTOLAKIS: And they were not Office of Research, were they?

MS. LOIS: It was Office of Research, but the people that are using it for event evaluation. So they were practitioners.

MS. COOPER: I thought that they had a contractor doing it. It was paid for by the Office of Research.

MS. LOIS: Yes, the analysis was done by a contractor. But the same contractor is helping the research to do an event evaluation, and including this analysis.

MR. BLEY: I think some of the driving force of having developers as part of the analysis teams was the last benchmark had people who didn't know the methods at all involved, and that sure didn't work very well.

MEMBER STETKAR: I'm just trying to get a sense for some of the comments that were made that it didn't appear as if some of the teams maybe spent as much effort as some of the other teams.

And one way of saying, okay, if they are only working part time, participating in what they thought was maybe just kind of an interesting research project, you might understand that, well, it's something that I have to submit before a certain calendar time.

But if they really were teams that represented more the champions of the methodology, they should have had anyway more of a vested interest to show that the methodology as applied by well-versed practitioners could develop reasonable results.

MR. BLEY: One indication of that was,

every – and this surprised me – all 14 analysis teams showed up at that workshop in October. I was very surprised.

MEMBER STETKAR: A lot of them really cared about doing well. There were methods they used, and they wanted the – so in that sense perhaps the differences in apparent level of detail in something you highlighted, the qualitative analysis, could be more attributed to the methodology than the analysts' team.

MR. JULIUS: No, I can confirm what John is saying. I know from the experience that my team that was doing the COSMIC decision tree, we were participating in the workshop, but we took some shortcuts when we did the analysis, and didn't do the full blown approach that we normally did, because we were late, and doing it late, and we had competing factors. And so we didn't go out and do things like, do the operator interviews or discuss – we used – drew on our bank of previous operator interviews, things like that.

So it is a valid comment.

MEMBER STETKAR: So Jeff, even though you are let's say a champion of that particular methodology –

MR. JULIUS: Yes, we are a champion of that particular method. We took some shortcuts, and we learned some lessons from that. So that is a valid comment, thought.

MR. FORESTER: We had one team that even did two different methods. They did a decision tree type of approach, and a CREAM approach.

MEMBER STETKAR: I happen to know those guys.

(Simultaneous speakers.)

MEMBER STETKAR: They took a lot of pride in their applications.

CHAIR APOSTOLAKIS: Are we done?

MS. LOIS: Yes, I think we are.

CHAIR APOSTOLAKIS: All right, maybe John, you can make a few comments?

MEMBER STETKAR: Yes.

I think my first comment is that I think that this is – what has been done and continuing with this process is extremely useful. I echo George's sentiment that it shouldn't be an open-ended research program, because it should have a goal in mind, it should have a well defined goal and some evidence that we are achieving that.

But the process itself of running

experiments; collecting information – and I won't call that data; just information – is extremely useful, and in a structured way that has I don't think ever been done is just a great great contribution to the field.

I think that – and it came out in the discussions today – the group should be more careful about the potential variability that is introduced into the analysis process by the way that the HFEs are defined, both in terms of success criteria, pass/fail, and in particular these time windows, because several of the methodologies do use time either as an explicit parameter, or certainly as a performance shaping factor however it's interpreted.

So the introduction of a somewhat artificial time window perhaps can unduly bias the results that you are seeing.

I'd also be very interested in seeing the types of evaluations that you are doing, the post mortem evaluations, through the integrated analysis of all of the HFEs for a particular event. I know you are going to look at each of the HFEs individually for the tube rupture cases. I'd be interested in looking at the entire scenario to see if there are kind of integrated lessons that you can learn from some methodologies versus another.

Because in some sense the way that the HFEs are defined are kind of arbitrarily sliced up in the sense of that scenario in particular.

That also might answer some of the questions that Said had regarding the evaluation of the crew performance in an integrated sense.

Is there – if we are seeing large crew-to-crew variability for the particular HFE of isolating – identifying and isolating the steam generator, is that same variability consistently generated by the same – across the set of crews through the whole scenario?

And that requires you to look across all of those four for the tube rupture case. So if you can keep that in mind as you are doing your evaluations, that might be important information to confirm the indication that there may be a high degree of crew-to-crew variability.

Or is it just simply an artifice of where we took the snapshot in this particular event.

I think that's all I have to say.

CHAIR APOSTOLAKIS: Okay, thank you.

I think I made most of my comments during the meeting. But one thing I wanted to reiterate is this business of being fully aware of where we are going with all this.

Is there a road map someplace that tells us we are going to finish this exercise, and this will give us this input, output that will be input to some other activity that will lead to something else, and eventually we will have this, what the SRM wants.

If there isn't, I strongly urge you to do that. This cannot be an open-ended activity. And again don't let the best be the enemy of the good enough. We have to go beyond what we are doing now as an Agency, and I would like to see such a roadmap with some discussion.

And if you guys want to draft one and then come back to the subcommittee and discuss it, I would be very happy to do it. I will be happy to comment on it without making a – writing a letter or anything.

But at some point though I think the full committee should be briefed on this and possibly write a letter on the way things are going.

Erasmia, do you have any idea when that might happen?

MS. LOIS: I thought that it would be an appropriate time when we finish the SGTR, the whole scenario, the evaluation of the whole scenario, so that we have a more integrated understanding of what's going on and how the study results are.

CHAIR APOSTOLAKIS: How long will that be?

MS. LOIS: And that is not going to happen before September. Actually we have –

CHAIR APOSTOLAKIS: September is fine.

MS. LOIS: – we have plans for having another workshop with the teams next November to discuss the results of the new analysis. So probably fall we'll be in a position to brief the subcommittee and also the full committee.

CHAIR APOSTOLAKIS: The thing that concerns me is, you said earlier that this particular exercise here will be completed in 2009.

MS. LOIS: And that entails the Halden and Paul Scherrer Institute resources.

CHAIR APOSTOLAKIS: Yes, but 2009 and then I'm thinking again in terms of the ultimate product. If there are two or three other things that have to be completed before we answer the SRM, and you know, the 2009 with the other stuff becomes 2014, I don't know whether the commission would look at it benignly.

I really think we ought to have that road map and start thinking about – by the way, Hossein, whenever the commission issues an SRM, there is some deadline there.

MR. NOURBAKHSH: We always send them – I

don't know how they close that SRM, but we already sent a response to that SRM.

CHAIR APOSTOLAKIS: And we told them what we are thinking. But at some point they will want to –

MR. NOURBAKHSH: At some point they are going to have to come back and tell us what –

CHAIR APOSTOLAKIS: What is the answer, you know, they can do that.

So –

MR. NOURBAKHSH: It may be even on the schedule or something that we have to answer. We have to check that.

CHAIR APOSTOLAKIS: Now the committee meets with the commissioners when, in June?

MR. NOURBAKHSH: June I believe, yes.

CHAIR APOSTOLAKIS: Is the HRA part of the –

MR. NOURBAKHSH: I don't think they have proposed items on that. It's up to us to propose items.

CHAIR APOSTOLAKIS: Anyway I don't want us to be caught by surprise and have a new SRM that says, you know, by June of `09 tell us what the results are.

So please, think about this roadmap, and maybe when you have something that is, whatever, halfway decent, we can have another subcommittee meeting, maybe half a day. But we really need that, to know where we are going, and how fast we should be going.

You can't go to the commission and say, we had useful insights. I mean that doesn't help anybody. That is a means objective to getting somewhere else which is concrete.

Anybody else who wants to make a comment?

The public? Is there a public?

(No audible response.)

CHAIR APOSTOLAKIS: Okay, thank you very much again. This is good work. This is good work. It has to be used widely.

MS. LOIS: Thank you.

CHAIR APOSTOLAKIS: So with that we are adjourned.

(Whereupon at 4:46 p.m. the proceeding in the above-entitled matter was adjourned)