Official Transcript of Proceedings

NUCLEAR REGULATORY COMMISSION

Title:	Advisory Committee on Reactor Safeguards Subcommittee on Human Factors
Docket Number:	(not applicable)
Location:	Rockville, Maryland
Date:	Tuesday, September 10, 2002
Work Order No.:	NRC-522 Pages 1-242

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

SUBCOMMITTEE ON HUMAN FACTORS

+ + + + +

TUESDAY,

SEPTEMBER 10, 2002

+ + + + +

The Subcommittee met at 8:30 a.m. in Room T2B3, Two White Flint North, Rockville, Maryland, Dana Powers, Chairman, presiding.

ACRS MEMBERS PRESENT:

DANA A. POWERS	Chairman
GEORGE APOSTOLAKIS	Member
MARIO V. BONACA	Member
F. PETER FORD	Member
THOMAS S. KRESS	Member
GRAHAM M. LEITCH	Member
STEPHEN L. ROSEN	Member
JOHN D. SIEBER	Member
GRAHAM B. WALLIS	Member

NRC STAFF PRESENT:

MEDHAT EL-ZEFTAWY	Designated Federal Official
AUGUST CRONENBERG	Cognizant Staff Engineer
MARK CUNNINGHAM	NRC Staff
JOHN FLACA	NRC Staff
ERASMIA LOIS	NRC Staff
SCOTT NEWBERRY	NRC Staff
J.J. PERSENSKY	NRC Staff
NATHAN SIU	NRC Staff

C-O-N-T-E-N-T-S

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

(8:47 a.m.)

MR. POWERS: The purpose of this subcommittee is for the staff to inform the ACRS on recent progress related to the agency's research programs on human reliability analysis and human factors.

I will caution you that the ACRS tends to glump this whole thing together as human factors or human performance. Sometimes that causes some confusion in nomenclature, so indulge us in our peculiar resistance to making fine distinctions in this area.

The purpose and the scope of these activities will be discussed as well as the relationship between the two disciplines. Presentations will include examples of how human factors, data, and information are incorporated into agency, human reliability tools, and how HRA can be used to identify and prioritize human factors data and research needs. Hopefully we'll discuss those research needs.

Gus Cronenberg is the cognizant staff engineer for the meeting and knows more about it than all the rest of us combined I'm sure. Medhat el-

Zeftawy is the designated federal official.

Rules for participation in today's meeting have been announced as part of the notice of this meeting previously published in the Federal Register of August 22, 2002. A transcript of the meeting is being kept. Open portions of this transcript will be made available as stated in the Federal Register Notice.

It is requested that speakers first identifying themselves and speak with sufficient clarity and volume so that they can be readily heard.

We have received no written comments or request for time to make oral statements from the members of the public for this meeting.

Before we get started here, I want to give the members just a little bit of background. The purpose of the meeting is to understand where the agency is going in its human factors research. Again, using the word "human factors" to cover human reliability, human performance, and anything else that has human involved in it.

The ACRS has been on record as recognizing that human factors is the emerging reactor safety issue of the future. On the other hand, ACRS has been relatively critical of many of the plans that the agency has put together to attempt to coordinate all the activities involving the word "human" within the agency.

Today we're going to be more focused, focused primarily on the research activities. And in developing this agenda with Dr. Siu, I thought that what we should concentrate on, it clearly would be useful to get the subcommittee educated on what has transpired since we've got together last time. But it's far more important for us to understand what the agency needs are, what the plans are to address those needs, and how well those tools, models, and understanding need to be developed in order to achieve what the agency needs to achieve in this area.

In fact, we've developed an agenda that allows copious time for discussion of what may seem philosophical issues. But I think it's important here that we have a good understanding of what the thinking is behind the strategy to not only understand what's going to be done but why it's going to be done and how well it's going to be done.

The intention is in fact to produce a letter to the Commission reporting what we have found about this human factors research program since it doesn't really mesh well with the plans for the research report itself. So, we're going to address it separately.

Consequently, I am going to poll the members twice today on what their thinking is. Once just before the break for lunch, which should pretty much bring to conclusion any of the formal presentations, and once after we have completed our discussions with the members of the staff in this area so that we have a good understanding of what our positions are and what our thinking in these subject is.

Do any other members have comments they want to make before we get started?

(No response.)

MR. POWERS: In that case, I'll call upon Scott Newberry to open up the proceedings here while Nathan sorts out whatever hat he's wearing today.

MR. NEWBERRY: Thank you, Mr. Chairman. I'm glad to be here. I wanted to come this morning and kick off the presentation and introduce the folks here at the table.

I think that you did a good job going through the objectives of the brief. That's our understanding of the, to discuss aspects of human reliability and human factors and all elements or interactions pertaining to those areas.

By way of introductions, of course, Nathan Siu, to my left, you all know. I want to mention that there's a bit of a transition going on in my staff. I'm bringing some work from Nathan to Erasmia Lois on my right, who will be giving a lot of the presentation today. So I'll just point that out to you. And of course, Jay Persensky to my right, who works for Farouk.

These programs are in two different divisions, which is also interesting I think, that human factors is under Farouk and the human reliability is in the risk assessment division and research. That's a topic that we revisit periodically in terms of whether that's best. So, this is a joint division brief.

MR. POWERS: I would just comment that it's been my perception that research as an institution here at NRC has been showing an enormous capacity to work across organizational lines. And I point to the PTS as an example of where that's been particularly effective. So I'm not sure that I would be apologetic about having things in two different organizations as an ipso facto sort of thing.

MR. NEWBERRY: Well, I don't want to come across as apologizing. I think we continually try to look at better ways to do business, not just communicate. But you do have a team approach here on this brief, which is I guess what I wanted to mentioned.

My remarks will be brief. I'm going to go through the objectives of the brief a little bit. I'll go through the outline of the brief and talk about some of the reasons we think this program is important. Then I'll excuse myself to head off to another brief.

But before I get into the briefing objectives and outline, I thought I'd mentioned two or three things. First, I hope you'll see today that we've been responsive to a previous input from the committee. You reviewed the research program last year, and we talked with you about that. We sent you a letter in terms of your comments on the methods development and where we should move the program. I hope you'll see that we've done that. You'll see a pretty extensive list of applications, PTS being one you mentioned Dana, where this work is important.

We've been trying to get to you but have been doing other things since 9-11. Some of the people here have been working hard since last

September. I think we wanted to get over here sooner but weren't able to do that.

The last thing I'll mention, and it was certainly emphasized in a recent SRM received from the Commission on our budget, and that is the need to constantly revisit our programs to see if they need to be altered, increased in scope or depth, or even sunset.

Even in the meeting with the committee yesterday on Reg 1174, the issue of David-Besse came up. It might come up today. I wouldn't be surprised if it came up, so I thought I would just indicate to the committee that in the context of our programs, and I think in this one, we are considering re-engaging the Commission on what should be done on the experience this year that could relate to safety culture research efforts. That would be the plan I would think, that we would have to re-engage the Commission given past guidance that they had given us before we set a direction. So, that's on our plate and I wanted to mention that up front before going into the view graphs.

Let's go to the objectives of the brief, which I don't think I have to spend much time on because the Chairman already mentioned them. But, we're going to provide an overview of the program, the activities in the program, and try to emphasize the relationships between human factors and human reliability aspects. Then, of course, we look forward to getting feedback from the committee. It's going to be an interactive discussion. That's what we've planned for.

Next slide. I won't read the view graph, but I'm going to go into a little bit of why we think these activities are important. I'm hopefully that you'll find Bruce Hallbert's presentation, a little bit later on the agenda, interesting and will provide some context for how the program overall relates human factors and human reliability work.

Next slide. There's considerable activity right now across the agency in terms of rule-making, licensing, the oversight process, and just the basic infrastructure itself in terms of where we prioritize what we think is important, etcetera. I think you'll see today that this program provides consider input to a number of those areas, PTS being one that Dr. Powers mentioned. But there's a broad need in my opinion across the agency for input from these programs.

PARTICIPANT: Could you eventually tell us what the specific useable outputs will be, which you'll be providing to these other --

MR. NEWBERRY: Yes, my point today is that it is our absolute intent to go through them.

PARTICIPANT: Useable outputs will be given to thermalhydraulics from this -

MR. NEWBERRY: I don't know that thermalhydraulics is going to be on the list, but you should see a matrix in my staff's discussion that you'll be able to engage on in detail.

MR. APOSTOLAKIS: Is it because thermalhydraulics is so fundamental it doesn't get any input from anything?

MR. POWERS: There's a major undertaking to understand why there are so many human errors committed in handling the momentum equation.

MR. NEWBERRY: In terms of operating experience, there are some major programs to learn from feedback. Certainly that's been the case this year. You'll see activities discussed today that get into all aspects in terms of the role of the operators certainly being able to provide recovery and prevent damage of the core, but also the possibility of worsening the situation.

Programs, the draw from our PRA experience, research programs, of course, line assessments done by the industry and work that we have done going back to things like IPE submittals and the like. But also, we're involved in reviewing proposals and applications from the industry.

I think one of things where I expected considerable time to be spent today is what's coming in the future, future trends, future events. I know the committee has been interested in interface issues, modifications to current control rooms, staffing policy, regulatory police involved with staffing as well as the new reactors coming down the pipe where there could be significant human factors/human reliability issues.

The agency is faced with a number of questions in terms of the impacts of these changes. From a regulatory point of view, certainly there's a question, I suppose quantitative sorts of questions that can be asked in terms of the impact on risk and how the human contribution to the risk profiles of plants manifest itself. And, we'll get into that a little bit today.

Let's go to the next view graph. I think Nathan pulled this together. It's really just a summary of what I mentioned to show that the human factors, PRA, or human reliability work -- providing input to the decisions that the agency is faced with.

PARTICIPANT: That's so general. It doesn't really tell me anything until you go into specific needs and specific outputs.

MR. NEWBERRY: Yes, it's very general. Sometimes it's not clear to some that our products are utilized in actual rule-making decisions, actual licensing decisions.

Just recently, I know Dr. Persensky and the staff provided a report to NRR that was requested and should be utilized in how to look at the monitoring aspect of the reactor oversight process in terms of looking at corrective action programs and the inspection program. So, that's what is meant by monitoring.

It was mentioned that we're doing work in the pressurized thermal shock area, which will come up today I'm sure. These folks are providing input to that integrated assessment of the current PTS rule. We'll have to see to what extent we should rely on the operator in the context of looking at potential modifications to that rule.

Then of course, the licensing decisions, where plants are ascribing to make a modification either going from a manual to an automatic feature or automatic to a manual feature. Those are licensing decisions, and we're working to provide input into that sort of decision.

Of course, all the way over to the left there are the agency performance goals, which we're trying to work towards. So, that's all the slide is trying to show in a general way. I know looking at my staff's view graphs, which you'll get into today, there is plenty of examples I think that would work from this outline.

Let's go to the next slide, just sort of a way of introduction, then I'll just move away from the table and let Erasmia and Jay take over the brief. I mentioned that Erasmia and Jay are the leads for the HRA and human factors research programs, and they'll be doing the brief today.

I think you've got copies of our programmatic material, which are referenced on the slide there in terms of the program plan, and the second paper, which outlines the human factors activity.

My interest in moving forward here as well, which I would mention, is not only to receive input from the committee but we're trying to give these plans a little bit more visibility. In both inside and outside the agency, I think we do need input. We need an understanding of where the work is being used. We're trying to do a better job at that, interfacing with the program offices, both NRR and NMSS. This is one step in that process.

I would suggest we go ahead and move ahead with the brief unless people have questions for me on my comments.

MR. POWERS: One of the issues that you may have touched on in your discussion was we tend to say the entirety of our human performance is focused on the performance of the licensees, and in fact, we have substantial activities within the agency itself where we have human performance most notably the inspection forces, both resident and nonresident at the various sites. Do I understand that you're thinking of looking into that aspect of human performance as well?

MR. PERSENSKY: If I may? I'm Jay Persensky.

One of the things that was in the second paper on the human factors aspects of the project was an attempt to transfer knowledge. I think that's the way I characterize it in that paper. The idea there was to develop some training programs for the staff, the inspection staff, that they had a better appreciation/understanding of some of the human factors issues as well as just recognition that it's time to call somebody else it. So, that's one of the topics that I have here as far as an infrastructure topic.

From the standpoint of nuclear power plants, from the materials side, we've actually been asked by NMSS to help them human factor, make their inspection modules easier to use. So, we're working with NMSS on that project right now. It's sort of a consultative effort as opposed to a major research effort, but we are providing some support in that area. We're moving in that direction slowly.

MR. POWERS: One of the big issues that's going to emerge tomorrow actually has to do with the ease with which the NRC staff can approach the significance determination process in the fire protection area. I mean it's a classic human performance kind of issue there. And so, I'm just asking are we thinking about human performance, not on the part of the licensee but on the part of the regulator now?

MR. PERSENSKY: The simple answer is "yes", we are.

MR. POWERS: Other questions for the now gone Mr. Newberry?

(Laughter.)

MR. NEWBERRY: I'm right here. I was just packing my bags.

MR. POWERS: If there are none, then let's proceed ahead.

MS. LOIS: My name is Erasmia Lois. I work for the Probabilistic Risk Analysis Branch of the Office of Research. I undertook recently the responsibility for the human reliability analysis program. We're in transition as Scott mentioned and Nathan had relayed before. He is here to answer your tough questions. I am going to do the easy ones.

Regarding background in HRA, I was involved earlier on at the NRC with the development of what we called in the early 90s predicted performance indicators through plant programs, program effectiveness, maintenance, training, etcetera. Then I moved on to review IPs and that gave me the opportunity to really comprehend the importance of HRA with respect to the PRAs. And recently, I've been involved in developing standards, PRA guidance. That also involves HRA.

Regarding the outline, I'm going to first address the relationships of human reliability factors then I will present an overall status of the plan, what we have right now, activities that are going on right now. Then I'm going to address a couple of specific activities, the advanced reactors, and the data collection and analysis project.

Next slide. This attempts to present the interfaces of the human reliability and human factors work. Human reliability is part of PRA, and PRA draws on many disciplines: nucleonics, thermohydraulics, etcetera. HRA is the part of PRA that helps model --- understanding of human performance under accident conditions.

The models, and they tell that we need to do a PRA, come from work that is done from human factors engineering and related disciplines: psychology, etcetera. So human factors is focusing on comprehending human performance in nuclear power plants and under accident conditions. Models and data developed there are used by HRA. Also, human factors work in research. They define new issues that we should cover as part of human reliability analysis.

As an output from performing HRA, we could provide or are providing to human factors work area that they may focus, they may need to focus more of their work scenarios or specific contexts.

HRA modeling needs, we have -- and also, how to help human factors work to prioritize their issues for work to be done.

MR. LEITCH: It seems to me, most of the current vintage of plants were built with digital instrument control systems -- I mean analog instrument control systems I should say. Many of the replacements are digital. Some of the replacements are being done piecemeal as the system is obsolete. There is a digital replacement for a particular compound.

Now I would think that whole issue of how that information is presented to the operator would be, as Dana says, something with "human" in it. But I'm trying to get clear, would that be something that was analyzed in the human factors or human reliability?

MR. PERSENSKY: It's primarily been a human factors effort to date. We'll be discussing some of that work. For the reasons that you just brought up, we are doing some work in that area.

MR. LEITCH: Because we have very little opportunity to design a completely new control room, but there are a number of modifications being made that influence operator performance.

MR. PERSENSKY: Right, and we're pretty much aware of those and we're tracking that both in terms of what we're doing here to develop review guidance. We're also with EPRI on their development of some guidance for the design of hybrid control rooms, which is what we call them.

MR. LEITCH: Okay. And you're going to get into that more later?

MR. PERSENSKY: Yes, I'll get into that later.

MR. LEITCH: Okay, thanks.

MS. LOIS: But also from my HRA perspective, as our comprehension and understanding is increased and the work is done at human factors, we plan to also improve our modeling capability and data capability for HRA analysis. So that's one feedback look. And, and I'm going to talk a little bit more later on that too.

MR. LEITCH: Okay, thank you.

MR. POWERS: The more I look at this slide, the more I like it because it has lots of things that can be the focus of our discussion.

One of those areas is the right side that says "PRA" and then it says "HRA". I think there's no question that the human reliability analysis that takes place in PRA presents a set of crucial questions, a set of crucial modes in there where you have to have probabilities that the operator will make an error of omission in the course of his activities. And, we put numbers in there.

What I struggle to understand are really two things. How well do we know those numbers that we put in there, and how well do we know the distribution of values that those numbers could actually adopt?

In the course of the day, I'd like to explore that to know better how well we know those numbers. If we know them well enough, that's one position. If we need to know them better then how do we go about knowing them better?

There have been a huge number of approaches for developing those numbers. I think I lost track right after the first one. But there's slim, odd and a whole bunch of things. Culminating perhaps in some Greek thing, which will forever remain nameless otherwise.

MR. APOSTOLAKIS: Misspelled too?

MR. POWERS: I don't know whether they misspelled it or whether the Greeks misspelled.

I'd like to have some understanding of

where we stand there. It boils down to the question of do we know things well enough there?

I'll comment that a source of confusion to the ACRS or surprise confusion on my part -- the rest of it was just surprise -- is that we've gotten a string of power uprates coming before the committee in which the times available to the operators to do things have been shorten. Of course, people looked at those and said does that have any impact on the safety and reliability?

In general, the conclusion from both the people applying for the license or the power extension was that "no", there was no real impact. The reviewers said the same thing. But, there was never any what I'd call a detailed analysis that said we've taken these variety of methods for estimating human reliability and the vast amount of data that we have available to supports those, we found that verily this was true.

We did get some interesting numbers in which relatively fine distinctions and probability were made that seemed to be contrary to our intuition on how accurately these HRA numbers can be estimated. So, any clarification you could provide in that area would be extraordinarily useful.

I do like the slide because it says that there is a feedback between human factors and the HRA models. And, I'd like to understand that better.

MR. FORD: On that issue -- and I'm new to this field so please excuse the simplicity of the question. HRA I understand, which is just the probability that such and such an action will take place at such and such a time.

What is human factors? Just how to improve on that reaction time and reliability? Is it ergonomics and things of this nature? Or, in that scenario, give an example of human factors?

MR. PERSENSKY: Well, as you said, the ergonomics, the timing -- human factors is a multi-disciplinary science or discipline. It's often referred to as human factors engineering. It's most commonly heard, if you listen to ads and things like that in terms of ergonomics. It addresses views and things like that.

From a more scientific standpoint, it gets into the issues of training procedures, the qualifications of the people that are doing the work, the man/machine interface. It's the whole picture of how the person interacts with the system.

MR. FORD: Okay, so it's a way of

improving on the actual data of HRA?

MR. PERSENSKY: One of the outputs in this figure here is that it would be in fact to help build a better database or to improve on the data that is used in the HRA models.

MR. FORD: Thank you.

MR. SIU: I'd just like to comment. HRA, certainly one of its functions is indeed to provide numbers that go into the PRA. But HRA also develops the, if you will, the input, the variables, the parameters. It defines those parameters. It says what are the errors that can occur or need to be considered?

So there's a qualitative aspect to that as well. There's an issue of what are the factors that affect the likelihood of those acts succeeding or failing. That's clearly where the --

MR. FORD: And the feedback is to somewhat control the input parameters to the HRA.

MR. SIU: That's right.

MS. LOIS: The example, the second half of this morning's presentation will help clarify that issue.

Regarding the overall plan status, as Scott mentioned before, we're behind because of unfortunately September 11^{th}. The last plan update is May of 2001. It's a five-year program. Some activities are near completion. For example, the PTS work and the work on quantification, including how do you address uncertainty.

Other activities are underway or planned. We expect to update it to keep the plan alive. Therefore, dates and milestones will be updated and projects will be added/deleted. For example, vulnerability assessment was not part of the program.

Also, work on HRA guidance and standards. We plan to have a higher level plan, to have a higher level plan activity description.

Next slide please.

MR. POWERS: Let me ask you. When you say a "higher level plan", it seems to me in the HRA area, it's more than just the numbers. It's the identification of where errors of omission can be made. That's inherently a qualitative thing. You just do that, and you do the best you can. People critique you and over time it gets refined. By now for existing reactors, I guess we've kind of got it. I don't know that that's the case, but my hope is it's the case.

But the numbers themselves, you put those numbers in and you say the probability of human error is 1 chance out of 100 this guy will make a mistake. And then somebody says well, how accurate is that? Is it 1 chance out of 100, or is it 1.1 chances out of 100? You snicker and say it's between 1 in 10 and 1 in 1,000, is that good enough? How do I know that's the case?

How do I persuade Dr. Ford over here, who only understands corrosion potentials, and insist -- I mean he can understand corrosion potentials because he can calculate them and then he can compare them against experimental data. And if the curve doesn't go through the lines, he does something to his model to calculate it better, right?

How do I do a corresponding thing over here to persuade him that the number I'm putting in there has some relationship to reality?

MS. LOIS: We hope that we'll address this question with demonstrating how we plan to collect some data that will provide more objective values in those estimates.

MR. WALLIS: I guess what the Chairman is getting at, is there some kind of an academic discipline or something?

MR. FORD: Is there an algorism to show

response time, frequency of response times?

MR. WALLIS: Or are you charting new territory all the time here, or is there some standard way of doing it, which is established and recognized and believable?

MS. LOIS: We have the opportunity through simulator exercises to kind of establish response time. I mean we get the time through thermohydraulics. And then how well people respond to that, the only real -- the best data we can have is through simulator experiments, and that's exactly what we're going to --

MR. FORD: But do you have a distribution of response times from the simulator experiments? Can you put down that that response time is an algorism of each of the operators or experiments of the operator?

MR. APOSTOLAKIS: No, you can't.

MR. SIU: At this point, we can't. As Erasmia is saying, we're trying to collect empirical data. That collection won't be to just go out and collect data, of course. There are qualitative models that say there is certain things that seem to be important that affect performance. In fact, you're going to hear a nice presentation on that later today.

What you'll see also of course is that we don't yet have the mechanistic model that takes us from these factors all the way to a human failure event in the PRA model, which can be lots of micro-errors and micro-recoveries all swished together into some general functional failure. I think that's something we could be driving towards.

I know Jay has been perusing some of these things. When we talk about simulation models for example for operators, one might hope to eventually develop that kind of mechanistic representation. We certainly don't have it at this point.

MR. POWERS: One of the topics that has come before the committee in just recent months in this regard has to do with the power uprates again. The particular issue, people assigned some probability of human error.   I think it was 1 in 100. When we asked the applicant "do you test on this in your simulator", he said "oh, yes. We test on it regularly."   "How quickly do the operators respond?" He said, "Within about 30 seconds." They never failed to do it correctly.

It was 52 times in one case that they had never failed. And in all cases, the response time was within 30 seconds. But they still used 1 in 100 as the failure probability. That seemed to be a complete mystery to everyone. I mean why that number in the face of all this empirical evidence?

And of course people said, "simulators are one thing, actual planned events are quite another." So, to account for that. But, that still didn't answer Dr. Wallis' question of why 1 in 100 and not 1 in 10?

MR. SIU: Maybe we should continue, but just a quick response on that, Dr. Powers.

Of course, one of the notions behind ATHENA was that you try to look for the conditions under which failure might occur, that might prompt the failure. Not knowing anything about the example you're talking about, I don't know how the conditions space was probed to see if they could challenge the operators in something that goes beyond --

MR. POWERS: They used THERP.

MR. SIU: Well, you're saying there's a certain set of empirical data but it covered a certain set of phase space, if you will. The question is are there other parts of phase space that might be risk important that were not probed and therefore, how do you deal with that?

I guess all I can say is that in things

like PTS what we're trying to do is to use evidence from talking with crews or trainers of crews and blend that in to say under this circumstance, how likely do you think success would be? But again, we don't have the mechanistic model for doing that.

MR. BONACA: One question I have is in your plan you talk about going to look at current symptom-oriented procedures. And that was a suggestion that we made about two years ago. Is there a plan already in place to do that?

I guess the feeling is that there is so much information there that could be very effective. Because I know for one -- I participated in some of them -- there is an enormous amount of information developed to build the outcomes of the procedures. And they're symptom-oriented in a sense. There was a lot of effort to determine the likelihood of the number of possible outcomes from a reaction. One would be more successful than the other would be.

So, I would like to hear more about the plan that you have to do that. I know you have it in your plan.

And also, the accessibility of this information to you. I mean will the licensee make it accessible to you? Is it available? I don't know if it's the right time to ask that question, but I would like to hear about that.

MR. SIU: At this point, quite honestly, we haven't done anything on that. We had put it in our plan. We had full intentions of doing work on that problem, but again, with other activities getting in the way, we just haven't gotten to it.

MR. BONACA: Because I wanted to say there were literally hundreds or many years of simulator data collected, reflected in those symptom-oriented procedures. I mean the BWR effort last years with iterations and iterations and refinements. So there is a huge volume of work there. And if it's accessible from the vendors, I think it would be a great help.

It's being collected under this program where you have a different kind of reaction and objective than the one that the simulator people were using at that time or the symptom-oriented people were using. So, I would really encourage you to get access to that information.

MR. APOSTOLAKIS: Coming to this slide, some questions I guess should be addressed to the slides these guys prepared.

It says SPAR models under the conventional reactors for monitoring. It seems to me you have the reactor oversight process on the left. It seems to me that you can help the NRC inspectors to do their job a little better.

It's still a question mark in my mind why there were no reports that I know of from the inspectors that things were happening that were out of the ordinary. The first reaction of course is to still blaming the utility, but it's not clear to me why the frequent change of various filters and so were not noted in some papers and notices.

So the SPAR models again -- the PAR out of course and so on -- but it seems it would be useful for this work to also address the issue of NRC inspectors. Is that going to be done?

MS. LOIS: We have that as part of the infrastructure, which addresses all of that.

MR. APOSTOLAKIS: Oh, okay.

MS. LOIS: It's actually embedded in guidance development.

MR. APOSTOLAKIS: Okay, because I was a bit misled by the word "SPAR models". Maybe you can put a few more words there. Or, maybe that's what you're doing right now?

MS. LOIS: What we have over here is kind of an analysis. Although it is not clear cut, these are analysis types of tasks. And, over here is guidance or standards development, which support those tasks as well as methods and tools.

Regarding the issue that you said, we plan to develop a guidance for the inspectors of the plan to help them identify human performance issues. That will come out events assessment as well as the experience we have through the PRAs and ATHENA applications.

MR. APOSTOLAKIS: Is this only HRA activities?

MS. LOIS: This is just HRA activities. Recently, the fitness-for-duty, our role is under revision and we were asked to provide a risk basis if possible. So that's one of the potentials. We haven't engaged anything on that. But these are activities that Nathan is pursuing, and I don't think we have concrete plans on that yet.

On waste and materials, we've completed some work for dry cask. We also communicated with NMSS and we frequently respond to questions.

On the advanced reactors, the plan includes the upgrade and advance as one element. I'm going to talk a little bit more about what we're going to do in this area.

MR. APOSTOLAKIS: The upgrade is the new INC?

MS. LOIS: The new INC, that's right.

MR. APOSTOLAKIS: Okay.

MS. LOIS: So then on the conventional side of the reactors, we are completing the PPS work, PRA, HRA. Also we have work on fire, steam generator, tube rupture. We haven't done anything yet, but it's in the plan.

MR. APOSTOLAKIS: What do you mean by that, the sequence? What happens in the accident sequence initiated by your tube rupture?

MS. LOIS: Yes. And do a more detailed PRA as part of that HRA.

MR. POWERS: My comment -- I was excited to see that because when this committee looked at the steam generator tube rupture accident in a fair amount of detail, we found a fully chaotic situation with respect to human reliability in obtaining flows of coolant into the system as the function of the number of tubes ruptured.

Surprisingly, they all came up with pretty much the same answer for the probabilities, but you didn't come away with saying, "Yes, that is the number." All you came away with was the feeling that human reliability and analysts talk to each other enough that they always come up with the same answer.

MS. LOIS: So that's an area that we're going to do work to probably come up with a better answer.

Aging cables is something that we're not quite sure if we'll do right now. There is preliminary work going on in that area. If the PRA is going to happen, HRA will be part of it.

MR. POWERS: Can you tell me what it means? I mean cable aging and human factors seem just about as orthogonal as -- I mean maybe they're not totally orthogonal. Humans age too.

(Laughter.)

MS. LOIS: Do you want to answer? Yes, go ahead.

MR. SIU: The issue here is that as the cables age their resistance to the environment is reduced. Now what are the cables on containment? A lot of cables are instrumentation cables. So the question is what would be a response of the operators if you have wide scale effects on instrumentation?

This is a relatively minor part of a larger activity. So what is showing are a number of applications to which HR is providing support. It's not necessarily a big program here.

MR. POWERS: So, you're going to look at procedures that the operators have and say if this particular device is producing spurious signals, erroneous signals, will the operator in fact be able to deduce that the device is no longer reliable, and can he then find other sources of information that give him the equivalent?

Is this not a topic that the licensees address a great deal of deal?

MR. LEITCH: There's a reg guide that describes post-accident instruments that will survive the accident. In most control rooms that I've been associated with, the instruments clearly annotated as to which instruments they are. The operators are trained to use that particular set of instruments in an accident situation.

MR. POWERS: Isn't it true that, especially in the emergency operating procedures, that the operators are enjoined to question their instruments and be skeptical of what they're providing at every juncture?

MR. LEITCH: Well, I think the general feeling is to believe the instruments. But when there's a discrepancy between the instruments, there's a preferred set of instruments that should be used and they're the ones that you should go by.

There maybe many indications of a particular parameter, and there's a set of instruments that are survivable through the accident and they're the ones that you're trained to go by.

MR. SIEBER: I think in general during emergencies, operators are told to trust your instruments but to crosscheck.

MR. POWERS: That's what I mean by skeptical.

MR. SIEBER: But the crosscheck is different than just saying this instrument if off scale high, and I don't believe it so I'm not going to do the action. That's not what they're taught.

MR. FORD: Could I ask a question?

MR. SIU: Sorry, I just wanted to follow up please if I may.

Again, I don't want to give the impression that the activities you see here are all development activities. Sometimes we're just being asked to provide support to say what is the risk significance of a particular issue. And the risk significance of course involves the human component as well as the hardware component. This particular project would also involve thermohydraulics, INC, and so forth.

This is simply indicating, as Scott indicated in the morning, we are doing a number of applications. This is one. Clearly, when we start digging into it, we would be looking at the guidance of the operators. Hopefully, we'll have the chance to talk to the training supervisors and so forth, and see what are indeed the expected reactions of crews under various situations.

MR. ROSEN: In Scott's introduction, he talked about the issue with Davis-Besse, and Dr. Apostolakis mentioned it also, and the need to think about safety and that sort of thing.

Part of that thinking leads me to a conclusion that we need some sort of early warning system on human performance and enhanced organizational performance. That organizational performance, which is the sum of all of the individual human performances, is degrading. And, I don't see any activities here that would lead me to the conclusion that this research is within the grapple with that.

That's just a question that's sitting here in front of me. I don't when you'd address that, but I certainly would like you to sometime today.

MR. FORD: I have a similar question. On the reliability analyses as I understand it, there is a lot of data for conventional reactors in terms of many years of information so you can come up with a distribution of a response time or whatever. However, we don't have the algorisms to relate that distribution to a factor like the age of the operator or whether he's right handed or left handed or whatever.

Given that fact that you've got no prediction capabilities, how do you come up with the reliability analysis for advanced reactors for which there is very little data, operational data? What is the process by which you can come up with that reliability analysis?

MS. LOIS: I guess the short question is that we start out like we started out for the conventional reactors. Where we lack experience, we try to come up with -- looking at the other types of activities that potentially simulate the data or the issues of an advanced reactor type.

But in actuality, what we're going to talk about after is actually work that was performed for advanced reactors. And therefore, human factors work has been done from the perspective of operator performance, all of the operator staffing, etcetera.

MR. FORD: So we're going to have a presentation on that very topic?

MR. PERSENSKY: There will be a presentation regarding a specific project that was done making certain assumptions about advanced reactors, primarily more the light water, passive reactors, not so much the modular reactors. But it's work that we had done several years ago, and that will be presented later on.

The other aspect of that is we look to wherever we can. What other industries might have similar situations? The chemical industry for instance has a lot of the same kind of continuous operations. So, if they have done work that we can find and try to translate that information into -- both from the human factors standpoint as well as the human reliability.

One of the big issues with advanced reactors of course is the modular reactors where you have one operator for several modules or a few number of operators. And I'll get into that a little bit later on.

MR. APOSTOLAKIS: It seems to me though there is a philosophical point that needs to be clarified.

There are no physical or chemical laws that govern what's happening here, so we can't really apply the same rules that we apply to materials or other physical sciences, natural sciences in terms of confirming a correlation with probability distribution and so on.

Rather, what we're trying to do here is produce probability distributions that reflect the communities' state of knowledge as to how likely these things are. These are not predicted models. This distribution has to be consistent with what we know about this thing and related things. And that's what Jay just referred to. There may be other industries where there are similar situations. So, what is their experience? Is it consistent with what we're saying?

MR. FORD: So you will assume that 1 in 100 operations will be a defective operation, and therefore, what is the impact on the operation advanced reactor?

MR. APOSTOLAKIS: Well, yes. But first of all, it's never 1 in 100. It's always a probability distribution. That's why it's not testable. I mean it's what we know. But what you're trying to do is make that distribution consistent with the totality our knowledge. So to ask for an experimental verification really is not the right question here.

You continually improve or change as your state of knowledge changes. And certainly, Davis-Besse was a major input to that. It has been and they will have to address it.

Another thing, for example, in several instances we have seen that the operators have taken actions that were very innovative. They acted in a very clever way. Brown's Ferry was one. We have made a conscious decision I believe not to include such events in our analysis, right?

Very rarely you will see that the operators do something that is not in the procedures and saves the situation. I haven't seen any PRA that says that. It's usually something that is dictated already or have been trained on.

But anyway, the philosophical issue is that they're trying to reflect not just the whole PRA business. What are the probability distributions that are consistent with what we know about this subject? For example, to put the probability of error as one in nine -- not in nine, nine in ten, is probably inconsistent with what we know about operator training and past incidents and so on. One in hundred, we don't know if it's consistent.

MR. WALLIS: Well George, I'm bothered by your saying there's no experimental verification. If there's no experimental verification, what kind of verification can there be?

MR. APOSTOLAKIS: The experience.

MR. WALLIS: Well, that's experimental.

MR. APOSTOLAKIS: But it's not in a traditional sense.

MR. FORD: What you're saying is you can never improve on 1 in 10. Then therefore, what's the role of human factors? If the guy is tired then presumably he's going to have a one in five chance of making the wrong decision.

MR. APOSTOLAKIS: But they take that into account.

MR. FORD: So you can improve?

MR. APOSTOLAKIS: Yes, as your knowledge improves. If you look at what we were doing 20 years ago, the THERP that somebody mentioned -- I think Dana did -- the first models relied exclusively on the available time. I mean if you go to the original report by Swain, he says six minutes after the alarm the probability of failure to do the right thing is this. Then there was a second generation where people went deeper into the context and what are the factors that may affect performance and so on.

I'm sure there'll be a third generation. Maybe they're working already on the third generation. But, this is how you evolve. You start with something very simple. At that time, people thought that the available time was the controlling factor. Now we know that it's an important factor but it's not the only one.

MR. BONACA: Well, the development of procedures was exactly one to improve performance because before it was based much more on simply contact information on the part of the operator. But now, it's really prescribed. There's a lot of study that tries to eliminate some of the judgmental portion associated with the response to the machines, and to simply guide the operator through proven or believed successful scenarios.

So, there is the component there that has come in. Of course, the training, there are elements that have reinforced or made the likelihood of success --

MR. APOSTOLAKIS: Yes, but human

reliability and human error is a relatively recent discipline. Human factors has been around longer. But human error analysis, I mean there's a very good book published in 1990 I believe by Professor Riesen. There have been other books since then, but we're talking about the last 20 years or so. Rasmussen presented his categorization maybe in the 80s, very recent.

MS. LOIS: Unless there is any questions on this slide --

MR. APOSTOLAKIS: I think that in light of what happened to Davis-Besse, you need a bullet there, not necessarily using the word "safety culture" unless you have masochistic tendencies.

(Laughter.)

MR. APOSTOLAKIS: Put something else like -- human errors that lead to initiating events, because most of the HRA work until now has been really human reliability analysis of human actions after the initiating event. If we've learned anything, it's that humans can actually cause an initiating event.

Find the right words and put them there, but I think that's a very important thing. It goes back to Mr. Rosen's comment too and I think the rest of the committee feels it. Because I just said, as our state of knowledge changes, our models change. And certainly what happened last March or April or whenever it was, was a major change in our state of knowledge, right?

MR. BONACA: Could you glump it under latent error?

MR. APOSTOLAKIS: I don't want to glump it. I want it to be exclusive with arrows and things.

MR. BONACA: It would be a type of --

MR. APOSTOLAKIS: No, because latent errors are just plain lying dormant. Here, I'm talking about things actually happening. So the latent errors may be contributors to that, but they're not --

MR. WALLIS: Sometime while we're talking about generalities, I'd like to have some idea of how you show that a model works. In all other fields of science I know about, you can concoct all kinds of theories. Eventually, there's a confrontation with reality and you have to say does it work? I don't know what you do to show when your models are working or not working.

MR. SIU: I think in the presentation of Bruce Hallbert gives later today, you'll see a partial answer to that. There's still some gaps that need to be filled.

MS. LOIS: Mark Cunningham, why don't you go ahead.

MR. CUNNINGHAM: To go back to the point from Professor Apostolakis that Scott alluded to early on in the presentation, where the issue of what's occurred, Davis-Besse and that type of thing, have raised issues about whether or not we should be including in this planning effort issues such as safety culture or some variant of that.

As the committee knows, we're under some constraints on our ability to do that. But like Scott said, we're reassessing whether or not we should go back to the Commission raise the issue again with the Commission about the importance of this and the need to do research on this.

MR. APOSTOLAKIS: But the initiation of imitating events though, you're not constrained.

MR. CUNNINGHAM: That's true.

MR. APOSTOLAKIS: But I think you're right. You really have to go back to the Commission.

MR. POWERS: If I could come up to the PTS item up there. You're providing input there that's been mentioned several times.

When the program team involved in PTS has spoken in front of the committee, they have emphasized the statistical rigor with which they will be doing their various phenomenal logical studies. Is there an equal constraint on you for rigor in the human reliability inputs that you provide to that PTS program, and if there is, how do you carry it out?

MR. SIU: As Professor Apostolakis indicated, what we're doing in PTS of course is developing the distributions for the human failure event probabilities. And that's essentially expert elicitation process. Then we propagate those distributions to the rest of the model just as you would as a matter of course.

Lacking the phenomenal logical mechanistic models and lacking experimental evidence for these particular scenarios and the general model to take experimental evidence and bring it in to this particular arena, that's where we are.

I think when Erasmia gets to her data slide, we'll talk a little bit about what we're trying to do to move towards a stronger technical basis for these things. I think personally, it will take time to get there, but there's certainly a desire to start doing that to make better use of experimental facilities.

The modeling efforts are frankly going on in other parts of the human factors community regarding performance of people under challenging situations.

MR. POWERS: I have no objection to expert elicitation process, especially in a field where I think Professor Apostolakis said quite correctly the distribution there that you're attempting to put down is not a measure of reality. It's a measure of this objective belief of a cross-section in the community.

So I'm wondering how do you go about getting -- I mean what community do you probe? Are you probing the regulatory community, the contract community, or the licensee community? Maybe the answer is "yes".

MR. SIU: Yes, but in PTS, as I'm sure the committee has been briefed, we paid special attention in talking with the trainers of the crews and with SROs so that there were people who had experience with these crews under situations that were relevant to PTS. We think we got the right folks providing input into this elicitation process.

MR. POWERS: Yes, but if I were a trainer of people, I would have a tendency to think my training is tremendous and wonderfully effective as to my abilities to persuade people do to the right thing would be relatively high.

In fact, one of the characteristics that we found is that any time we elicit experts, they have a great deal or more confidence in their knowledge than probably is warranted.

MR. SIU: Yes. And what we tried to do, again not knowing what the underlying truth is, what we tried to do is make the people involved aware of this biases upfront. We tried to probe to again see what are the conditions that would lead you to a different performance level, how likely do you think those conditions might arise, bring in examples of how things have that happened in other situations and can that arise in this situation.

I think the belief of the team -- and John, you can add anything if you want -- John Forester of the PTS team. I think the belief was that we got some good input from them. They started thinking about these different situations. It still might be biased, but I think we've tried to address it was best we could.

John, do you want to add anything to that?

MR. FORESTER: I'm John Forester of Sandia National Labs. As Nathan said, I am on the PTS team and participated in the HRA.

In terms of the team that we tried to elicit to help us with the quantification process, particularly in the case of one of the plants at Palisades, we had not just trainers. We had people from operations. We had someone that went procedures, procedure development. We also had members of the HRA team: myself, Dennis Bley, and Alan Kozlowski.

All of us participated in the quantification process. You had a wide range of people. The idea is everybody brings information to the table, ideas that they have and their knowledge about how the scenario will evolve, what information will be relevant, what kind of things that might happen that could lead to confusion for the operators in actually performing the task.

So, the emphasis is on obtaining as wide a range of information as we can in performing the expert elicitation process.

In terms of biases, we try to control for biases. We try to use a facilitator, someone that leads the discussion to where there are possible biases and tries to correct for those and make people aware of the potential for them.

MS. LOIS: And that includes the simulator observations?

MR. FORESTER: Correct, we did do simulator observations. We watched the crews in related scenarios to see how they would perform.

MR. APOSTOLAKIS: I think facilitators are funny people frankly if you ask me.

(Laughter.)

MR. WALLIS: . . . if you think about Davis-Besse. If you asked Graham Leitch or people with experience with reactors to think about it before it had happened, could this sort of thing happen in the plant? They'd probably say they couldn't believe it would happen like that. It never happened in my plant.

So you're asking all these experts, and they would say the probability, this is extraordinarily small. Some kinds of conditions are there in that plant which made it happen.

MR. APOSTOLAKIS: That's one of the biases that John mentioned.

MR. WALLIS: So how do you do that?

MR. APOSTOLAKIS: There is nothing you can do. I mean you try. If the whole expert community is wrong, I really don't know what it is that you can do.

(Laughter.)

MR. POWERS: I get the impression that, I mean the sense, the note I've taken here is you do the best you can.

I will comment that in my own experience in doing these elicitations, particularly of operators of plants, not power plants but in fact research reactors, is that their answer to a particular question: could this ever have happened, is "not in my plant."

But look at these guys over in Idaho. Those guys can have this problem but not me. Those guys can. Of course, Idaho gives you exactly the same answers.

That in itself is a surprisingly common comment. In fact, I can't think in any of these issues where we were polling operators at energy and defense programs plants where we didn't get that response. "It won't happen here because we're very careful." But those guys, go talk to them. Go look at what they've got.

MR. APOSTOLAKIS: The truth of the matter is that before the three mile island accident, putting these operator errors in the PRA was a struggle.

MR. POWERS: Oh, yes.

MR. APOSTOLAKIS: Because the sponsor, the utility sponsor would tell you, "That can't happen in my plant." I mean that was a standard response. Things changed after three mile island.

But coming back to what this represents, I think it's important to make it clear -- you mentioned the expert community. Of course, expert community can mean a lot of things. But I think eventually your distributions here will reflect the state of knowledge of the experts in the human reliability area, at least in the United States but also broad because you participate in -- in fact, next week there's a major meeting that I understand you guys are going in full force. So, this is really what this is intended to represent, not just the views of Dr. Lois and Persensky and Dr. Siu.

Now there is always a reaction like you didn't use my model so this can't be any good. But at least they're not going to say, "Boy, your distribution is way off." It could be up by a factor of ten or something.

This is the same thing we're trying to do in the seismic area and in all cases where there are very rare events. You're really trying to capture the state of knowledge of the community, the entire community.

MR. ROSEN: I would like to comment on some of the views expressed here that the Davis-Besse situation would not have been predicted by those of us who have some knowledge of plant operations. I think that's incorrect.

I think with the data that's available or that will become available, had that data been put in front of Graham Leitch, Mario Bonaca, or maybe myself -- and I'm talking about things like the corrective action system performance and some other information perhaps out of the safety conscious work environment area. If that data had been visible or was visible to persons or a person who had a lot of experience, he could have predicted that the plant would have trouble, serious problems in the future -- not that the head would crack and the different things that we now know happened that would happen.

The culture was degrading, and serious and significant issues would rise at this plant in due time.

MR. APOSTOLAKIS: But still, I think one of the points that others have made -- and I agree with you on this. But suppose now you are a member of the group of experts that are helping Sandia and Idaho before Davis-Besse, and some crazy guy says, "You know, there may be a situation in the future where they will have multiple warnings of things that are going wrong and they will ignore them." Their corrective action program will not include hazard analysis and this and that.

Would that be a reasonable thing for somebody to say or would it be shut down by people who would say, "Our plants are not run that way."

MR. POWERS: I guarantee it would be formidable --

MR. APOSTOLAKIS: That's the risk that you will not think of unusual and very rare conditions. Given the conditions, I think it's pretty straightforward. So that's what I think John Forester was referring to. Experts can be wrong.

MR. ROSEN: There's no question that the scenario outlined led to a conclusion by someone that this plant was heading for trouble and that the plant managers and the rest would say, "No, that's not true. You're wrong." There's no question in my mind that that conclusion would be thought. But, that doesn't make the conclusion wrong. The very people who are fighting are the ones who are creating the problem.

MR. POWERS: That's right. Erasmia, as

Professor Apostolakis would say, "You're going so slowly."

(Laughter.)

MS. LOIS: So, I guess I want to --

MR. APOSTOLAKIS: I think you should say, "Next slide, please."

(Laughter.)

MS. LOIS: Next slide, please. Thank you.

This is an outline, a very broad outline of what we plan to do for advanced reactors and upgraded reactors. The objective is to determine if any improvements are needed to incorporate the influence in human performance in the PRAs for upgraded or advanced reactors.

The issues are the ones from the committee: reduced staff, changing the role of the operator, new control room design, multiple modules, and long-term recovery available for the accidents. What we are hopefully going to get out of it is what issues should we address, develop methods and tools to address those issues --

MR. POWERS: Can you articulate what you mean by "develop methods and tools" with any specificity at this point?

MS. LOIS: Probably not. If you look at,

for example, reduced staff, the HRA now has some underlying hypothesis as to how many operators are there, etcetera. So, we're going to look at the new proposed designs and their proposed staff in combination with potential accident scenarios and see how that plays out and changes the underlying hypothesis or even modeling in the HRA.

Do you want to add something to it?

MR. PERSENSKY: Yes, I'd like to add something. I think this is an opportunity where we're going to have a close cooperation. I've just initiated some work.

Nathan mentioned earlier that there are some techniques out there for behavioral modeling, how to model people's behavior, that have been applied in many military settings, particularly the Navy and the downsizing of their ships, especially the DV-21.

We're trying to see if we can adopt those models for use in the nuclear industry, particularly for this kind of thing where we really don't know yet, but we know that there's going to be some changes in the role. It's a function and task analysis based approach. That's the kind of model where we can feed in on this issue of reduced staff into the HRA model.

MR. POWERS: I guess I'm familiar

particularly with some of the Navy work because they have this problem too. You know, how many people have been put on a bridge, especially when you've got a highly instrumented and highly digitized bridge. People are expensive so you want to minimize those and still have proper coverage and things like. I mean they worry about these sorts of things.

But there's another approach that has intrigued me. I don't know whether the NRC gets involved in this. I know that MIT is involved in this. And that is these fairly fundamental -- I think you call them flatland kind of models, where they're trying to look at how social beings interact in a simulation sense.

Cooperative and competitive things have been most of the focus, but I've often wondered if those techniques don't have a place to play in these staffing issues. I just wondered if you have any contact with that or -- I mean it's highly simplified sort of thing. It probably is better for predicting how amebas work together right now. But it certainly yields some insights, certainly in the area of competitive and cooperative behavior.

MR. PERSENSKY: We're looking, and we're trying to keep abreast of that literature at this point. I know that there's a lot of work being done by DARPA and the Navy and the military in terms of how people interact. It's a lot of team interactions, joint decision-making. In fact, some of that's going to be presented at the conference that George had mentioned next week. So, it's work that DARPA is doing.

MR. POWERS: Yes. That's good.

MR. FORD: Could I ask a question on this? Given that some of the advanced reactor designs are somewhat conceptual right now, you don't know quantitatively the answers to the "what if" questions. Such as, if there's an accident scenario, you don't know what the operator reaction times would have to be in order to mitigate a series of actions.

How is your timing for this particular project, developing the methods and tools? What is the timing since you don't know what the target is?

MS. LOIS: I guess we're going to start out with existing designs that are better. For example, AP 600 and AP 1000, these are similar reactors in the sense that they do have the slow evolution of events, long recovery times.

Then based on probably simulator data as we discussed before -- PRA usually starts at a very high level and then as you gain knowledge, you go along and you improve your details.

MR. APOSTOLAKIS: I don't think they're going to produce distributions for advanced reactors. I think they're getting ready to address the issue later. For example, as Erasmia just said, now you're going to have to deal with very long operator response times, not just a few minutes.

So, you have to think about it. Are there existing models capable of doing this? Are there any additional factors I should include in the model, without necessarily saying for this particular advanced reactor, the fast reactor, this is the time and this is what I have to do.

MR. FORD: I guess my question is coming more as a research manager. You're asking -- I've got these conceptual designs coming along. I'll assume a worst case scenario that I'm going to have real slow operators and very few of them. As a research manager, how much money am I going to invest in developing what method, what tool to do what, to be improved on what?

MR. APOSTOLAKIS: I would phrase it a little differently. I have these new designs. Do they create any new context that I have not analyzed?

Then, the additional. Do they have any new dimensions to the problem that the existing models don't have?

MR. FORD: Well, you've mentioned AP 600 and AP 1000.

MR. APOSTOLAKIS: Well, AP 600 is really evolutionary.

MR. FORD: So what in the current tool box do you have for HRA that needs to be improved?

MR. APOSTOLAKIS: Yes.

MR. FORD: No, that's a question.

MR. APOSTOLAKIS: Oh, that's a question?

MR. FORD: Yes.

MR. APOSTOLAKIS: One of the things as I mentioned is nearly complete automation. I mean I don't know if it's there but --

MS. LOIS: The changing of the role of the operators.

MR. APOSTOLAKIS: Yes, the changing of the role of the operators.

MS. LOIS: So you might have just one guy watching over 10 models, one of two guys. That aspect of it --

MR. ROSEN: Erasmia, I have a problem with that. I think there is an irreducible minimum below which one cannot go in running nuclear power plants.

That is because it is not just that the operators sit there waiting to do something in the event of an accident. They're involved continuously in such activities as work control and authorization and wondering what's going on in the plant. People are out there doing things and there is a tremendous amount of communication coming up from the plant.

Also, in many plants they form the fire brigade around the clock. While we're sleeping or watching a ballgame, they are there in case there's a fire. They're the first responders.

So I think there's an irreducible number of operators no matter how much automation you --

MS. LOIS: Oh, I'm sorry for mentioning it.

MR. ROSEN: Maybe this is just a general comment because I don't believe these numbers.

MR. APOSTOLAKIS: I think the automation affects more the information that reaches the operators. I don't think the major issue is the number of -- because we don't really understand, as far as I know, the complete spectrum of failure modes of digitalizing.

MR. FORD: Doesn't the design for passive plant response -- like we see a lot of people advocating -- put you in the position that you've really got to confront the error of omission issue?

MR. POWERS: Yes. It seems to me that I would just highlight that. I've waited as long as I can. Now I've got to go attack the error of omission issue. It's been out there at least through last year's report.

MR. WALLIS: Let me bring you back to something that we've been already which is approving upgrades to power. There have been PRAs submitted, and we have had some things to say about those PRAs. What they have really come down to is simply saying the operators have more or less time to do certain things. Someone has made some estimates in those PRAs.

Do you folks think that those approaches were good? Were they adequate? How should we take those assessments which have already been submitted? What should we do to do it better? I think we'd like advice from you about that. This is going on. It's happened already and it's going to happen next month and so on.

MS. LOIS: The HRA plan suffers from initiating work and --

MR. WALLIS: You can't help us with any of those things?

MS. LOIS: Eventually.

MR. APOSTOLAKIS: The real question is how do you change the probability distributions when the available time changes.

MR. WALLIS: I'm nervous about that. I'm listening to the conversations and my colleagues are telling me they've got other things to do.

In my experience -- nothing to do with reactors but in a kitchen or something like that -- the more time I have, the more likely I am to make a mistake because something else intervenes. I've got to do this or that. I know I've got to this and I know I've got to do it in a minute, so I do it. If I've got five minutes, I say I've got five minutes and then something else happens, and it distracts me from this thing I've got to do in five minutes. I don't have time when other things are going on. But this is just interjection --

MR. APOSTOLAKIS: I think that's a good point. The sensitivity -- if you really want to look at those reactions like Dr. Wallis just said, this is happening now. We are approving power upgrade. And, the sensitivity of the human error probabilities to the available time is something that is of extreme interest. Maybe you can make them, but by the time you're done though, probably all the reactors will be operators.

MR. WALLIS: Well, at least you can look at it and give us some advice, right?

MR. APOSTOLAKIS: Yes.

MR. WALLIS: You're the experts we can turn to.

MR. FLACA: This is John Flaca. That's a good point. We have a synergism as to the activity that's going on. It's looking at all the changes that are going on in the outside world. One of these of course if power upgrade. In that context, I think that is an important issue to look at. And, I think we'll take that back with us.

MR. LOIS: Next slide please.

MR. WALLIS: Well besides looking at it, could you at least give us definite advice when you look at what's happening with power upgrades and when you look at the PRAs? Would somebody who knows in the agency make a decision about whether what they're doing is reasonable or not?

MR. APOSTOLAKIS: Nathan, you mentioned that -- was it Nathan or was it Scott? I don't remember -- that EPRI is involved in some of your work. Have you guys had a chance to look at their human reliability models? Do you use them?

MR. SIU: We haven't formally reviewed them. We had some interactions with them. I think you participated in that workshop we had here back in, last year I think it was, where they made a presentation on it. We know they've made progress since then. But, we haven't, "no.".

MR. ROSEN: One thing, as long as you've brought it up, EPRI as the leading indicator program. Are you aware of what they're doing there? This, to me, is a very exciting new approach. It may in fact lead to some visibility of the degradation in the future of plant operations because it gives you some insight into the safety culture.

It's basically a program that uses observational techniques to look at performance in the field, and each of the observers rates the operation as to whether it was good or not so good, whatever. The compilation of all this data ultimately can lead to some insight into whether the performance is improving, staying the same, or getting worse.

I have spoken to EPRI who are involved in that, and I know some utility people too, who would be willing -- and by the way, I've mentioned this to the Chairman of the subcommittee that these people would be willing to brief the ACRS at some point if we're interested.

MR. PERSENSKY: I'll add to that. I'll jump in here. I do work with the EPRI in the Performance Technology Subcommittee, and I've talked with them about the possibility of them coming in and taking with this subcommittee, not the whole ACRS, about the work they are doing in this area. They are willing to come. They do have a broad range of topics that you might be interested in.

MR. ROSEN: And in particular, to answer Dr. Apostolakis' question about their modeling, not just the leading indicator database and what's being done in the industry with that, but also the model of human performance and how it's used, I think I think there's one member of this subcommittee that would be interested now.

MS. LOIS: Next slide please.

MR. APOSTOLAKIS: Good idea.

(Laughter.)

MS. LOIS: Finally we get to the data collection and analysis. The objectives of that project is to determine the data needs for HRA, collect and analyze --

MR. WALLIS: I thought that George told us we couldn't do experiments. What is data?

MS. LOIS: It's existing information. That could be inspection reports, event reports --

MR. WALLIS: Is it word by mouth type of information or is it --

MS. LOIS: Documented information.

MR. PERSENSKY: Some of it might be simulator data that you might consider to be part of an experiment.

MR. ROSEN: I think, exactly. I think the idea that we don't have any human performance data is just wrong. Whether it's exactly applicable to the actual circumstances of a reactor one can argue, but we have lots of simulated data on whether operators take the prescribed actions within the symptom-oriented emergency operating procedures. And, that is valuable data.

MR. WALLIS: We have reams and reams of data.

MS. LOIS: Yes, that's one resource of data.

MR. WALLIS: Now I understand you're going to tell us more about how you're using that later, as

I understand?

MS. LOIS: Yes, yes. And therefore, the intent here is to really utilize and capitalize on as much as possible on existing information.

The work is to be performed to Idaho. It's co-funded by both programs, human factors and HRA. It currently focuses on the quantification aspects of it, ATHENA applications, which is by Sandia. Interfaces with international committees, CSNI has an effort on data collection and analysis. And also, the work supports Halden. It works with the Halden project.

MR. WALLIS: Go back to the number two bullet: collect and analyze data to support HRA model development and quantification. Is there some idea of the state of the art? I mean models have been developed, and I'm told there is a lot of data. Why aren't the present models good enough?

I have no idea from your discussion as to what sort of the state of the art of this field is in terms of what the models are. Questions that were asked at the beginning, how good are these numbers? I still don't have a good feel for that.

MR. SIU: Yes, and I think that goes back to, I think, Steve Rosen's point. We have information. The question is: Is that information applicable for the specific human failure events that we're looking at.

We looking at, which we all acknowledge, fairly rare conditions, very challenging. Generally, risk significance sequences. You failed a number of pieces of equipment and how do the operators respond to those particular conditions.

So, there is a question of applicability. There are also questions of if I vary certain factors, if I make changes to some of the things that maybe we'll get in to. Jay has an activity on fatigue. How do potential change and how we deal with fatigue in a regulatory space affect the risk profile? So you need models to be able to say what's that affect, and we don't have those at this point.

So, it's looking at not only the baseline numbers but the affects of those changes.

MR. WALLIS: You're saying all the things we don't have. Maybe it would help, and maybe it's been done before and I just missed it somewhere -- you actually had some demonstration that some model is useful and that some model represents some data.

MR. APOSTOLAKIS: I think in answer of your question Graham, about why aren't the current models good enough, first of all you have to appreciate that very, very few organizations in the world can afford by the NRC is doing here. They don't have national laboratory support and a lot of experts coming in.

What you see are models in the literature that tend to emphasis certain things that others don't emphasize. For example, some models from Europe tend to rely a lot on the centerpieces, the decision-making process in the minds of the operator. Then they ask themselves how is this affected by this and that.

Other models we've mentioned already tend to give a lot of emphasis to the available time for action. Other models do something else. You have models from Norway, from Sweden, from everywhere. But nobody has really spent the time and resources like these guys are doing to try to bring everything together.

MR. WALLIS: Models are fantasies until you can compare them with data.

MR. APOSTOLAKIS: That's right.

MR. WALLIS: It must have been done otherwise --

MR. FORD: As I understand it, we're going to see that this afternoon. We're going to see curves and data.

MR. SIU: This morning.

(Laughter.)

MR. FORD: So, your question may be answered.

MR. APOSTOLAKIS: But the basic approach of a physical scientist doesn't apply here.

MR. WALLIS: Yes, but something does.

MR. APOSTOLAKIS: You're dealing with a --

MR. FORD: But if you remember in the steam generator program, you saw distribution curves of a probability of detection.

MR. APOSTOLAKIS: Yes.

MR. FORD: And we had different curves for different teams, the good team and the bad team.

MR. APOSTOLAKIS: Right.

MR. FORD: Now, there's got to be a reason as to why the good team is good. Because of experiments or something like this.

MR. APOSTOLAKIS: But they make those distinctions too.

MR. FORD: Well, I think that's what Graham are struggling with. Let's see some data to back up these good and bad models.

MS. LOIS: The approach is to characterize information that is needed for HRA methods. We hope, as Dr. Apostolakis was mentioning before, we'll look at each one of the HRA methods available right now and identify what are the underlying hypothesis for the method to what types of data are needed.

And we're going to do that in a couple of steps that I have here. First, identify the concepts and terms used in the methods then identify the commonalities in the concepts. That will allow us to look at the data sources and mind them in a more systematic way as opposed to this particular method or that particular method.

Then we'll identify and evaluate data sources. And, we've done some of that work already. Then develop methods to use the data. Eventually, develop a method for estimating human error probability on the basis of the work done on the data collection.

Next slide please.

MR. POWERS: I guess one of the crucial questions that we really need to understand, there are a plethora of acronym methods for doing human reliability analysis, and that slide seems to say, I'm going to develop yet another one of those methods.

The question that we really have to understand is what is it that -- what need are you satisfying that these other things don't satisfy, and how accurately do you have to satisfy those needs?

MS. LOIS: That is part of guidance development that I had on slide before. We're going to address and examine each one of the available methods right now and provide guidance as to what are the characteristics of the method, what applications are appropriate, to what extent, what is the level, and potentially examine different applications, regular applications, and determine what is the level of detail or analysis needed, and therefore indicate what methods would satisfy that analysis.

MR. APOSTOLAKIS: Will you tell us at some point why CREAM, which is one of the models, is not good enough for the NRC? It has already been developed. Why the MARMUS model is not good enough for the NRC? I think that was the question.

Those guys have invested a lot of money. They have developed a model, and here we are developing another one. Why don't we just take the --

MR. SIU: If I can, I don't the point of Erasmia's slide is to say we're developing another method. What we're trying to say, and maybe we're not doing a very good job, is that there are a lot of sources of information out there. There are lots of sources of data. Sometimes these data are compiled by folks with a particular method in mind. So of course, they categorize information in a way, and collect it for that matter, the information to satisfy the needs of that model.

We need to be able to work with these folks to take the information they've got and make it useful in the activities that we've got going on. It may be along that along the way we find out that indeed there are some aspects of CREAM that we really do need to adopt in our approach or maybe it's in the MARMUS. I don't know that we've really thought along those lines yet. But, this is really an attempt to identify potentially useful sources of data and start making them available.

The notion of coming up with common technology is just a way that will help us communicate across all the various groups. I think there's a general recognition in the HRA, in this research community, that there is this plethora of methods and that we really do need to be working more closely together. And as part of this meeting coming up next week, we are going to be engaging with folks at CSNI to do work along these lines.

So again, we're not trying to say that we're going to create another method. The other thing I'd like to say is that there are a lot of commonalities. We talk about this long list of methods, but they have quite a bit of similarity.

MR. APOSTOLAKIS: I understand that the quantification effort is near completion. You did that using some sort of a model?

MR. LOIS: Some sort of what?

MR. APOSTOLAKIS: A model? Because Nathan just said you haven't yet looked at the other models and see what else they have that you may want to use. So, how does that --

MR. SIU: The quantification is really referring to bringing ATHENA to closure.

MR. APOSTOLAKIS: So it's not a model?

MR. SIU: No, it's an elicitation process. This is what was used in PTS.

MR. APOSTOLAKIS: I see.

MR. SIU: And that's where we are right now.

MR. APOSTOLAKIS: Okay. But I really want to emphasize that you really should do this. I mean before you embark on many developments, you should have a good evaluation of existing models with their advantages and disadvantages, merits and demerits. If the French have got something that is useful, you just go ahead and use it. If the Norwegians do it, fine.

This has been one major problem with this community. Every guy develops his own model, ignoring everybody else. This cannot go on.

MS. LOIS: But let me ask you something. Would you adopt a methodology that has been produced somewhere without having the capability to view it by actually seeing it, seeing the actual data that's created?

MR. APOSTOLAKIS: I said evaluate. Evaluate is all done. But, don't ignore it. Don't have an introduction that says oh, by the way, the following references also deal with this subject, 1 through 35. No. You say, CREAM has these good qualities and we're going to use them.

MR. SIU: We completely agree.

MR. APOSTOLAKIS: Very good.

MR. POWERS: When can we anticipate that we'll have this listing of 1 through 35, and here are the good features and here are the bad features?

MR. APOSTOLAKIS: At some point, we should.

MR. SIU: Yes. Again, we've been dancing all around Erasmia's presentation, but for every program or project that she has on that chart where we are talking about development needs -- and obviously we have to do that -- the applications, we're using the applications we've got in hand.

MR. APOSTOLAKIS: By the way, when I said this community, I was talking to a friend of mine who is in reactor physics and he told me there is nothing surprising about having some models. In the early days, when the guys were working in electronics, every organization in the country had its own transient code and this and that. Finally, things converged to something that's widely acceptable.

So even in the natural sciences, they things can happen. But, it's time to bring everything to closure.

MR. SIU: And again, I think we are actually trying to drive towards that closure.

MR. APOSTOLAKIS: Good.

MS. LOIS: This is the last slide. Then I conclude by presentation by mentioning again that the data generated for the advanced reactor staffing study will be discussed in some detail today. The objective of that discussion is to show collaboration of the two programs and how we can use existing information or create new information through simulator experiments.

MR. WALLIS: So what would your output? Is it going to replace this expert elicitation approach or what? What's going to be the results of this?

MR. SIU: I think in a long-term vision, that would really be nice. Whether we can get there, we'll have to see.

As we go through the presentation, as I said, you'll see some nice work that leads up to a point. But that point isn't necessarily the input to the HRA. There's a gap there, and we need to be able to address that gap. So, there's some technical work that needs to be done.

I think that we would certainly like to drive towards a more data based or at least data informed analysis. That's the vision of what we're trying to put forth. That's why we've put the data task as one of our top tasks in the program.

MR. APOSTOLAKIS: Have you found your collaboration with CSNI useful?

(laughter.)

MR. APOSTOLAKIS: I mean for a few years

now, I see that -- and I'll give you an example when I'm saying this. The NRC doesn't have it's own program and organizational factors, but we are in consistent conflict with our colleagues in Europe through CSNI.

Finally, I saw a paper from one of the countries. And, if you guys ever dare come here with a ridiculous piece of nonsense like that, this committee will probably not be kind to you.

MR. POWERS: It is my usual practice at this point to ask if there are any additional questions of this speaker. I think I know the answer to that, so I propose that we take a break until twenty of and then proceed with the rest of the presentations.

We can come back because I think there's a thought provoking presentation, certainly succinct in its visual aids that provoke a lot of questions.

(Whereupon, the committee recessed for a break from 10:22 a.m - 10:32 a.m.).

MR. POWERS: We'll begin by indulging the Chairman, who was reminded of a question by one of the audience that he failed to bring up. We had on the previous presentation quite a list of applications of HRA that are going on within the agency. John Flaca mentioned synergisms with some elaboration.

There is another area that is under active consideration by the agency and that is changing the categorization of equipment through the plant, retaining the functional requirements but not necessarily the elaborate QA and QC requirements that are placed on that equipment.

That equipment of course gets used by operators, and there must be some impact if not in the actual liability of the equipment, in the operators' perception of the reliability of that equipment. That should, in some sense, affect the human performance error rate associated with that equipment.

I didn't see any reference to application of HRA to those questions. I wondered if that was just because I didn't understand what synergism meant in its entirety or it's an omission or what the situation is.

MR. SIU: I think what we were trying to do with the guidance and standards bullet way at the bottom of Erasmia's chart, we need to provide information tools to users, let's say reviewers of applications to allow them to take advantage of HRA lessons without necessarily having to do an HRA.

We don't have an element that talks

specifically to let's say changes in reliability of equipment and how that might affect operator performance other than if we were doing a study in terms of context. But I don't know that we would be especially well tuned to get to that. So I guess that's one place where you could say we don't have something specific.

MR. POWERS: It seems to me that the ACRS, in its deliberations in connection with Option 2, has at various times made suggestions about the information communicated to the expert panels that should occupy the expert panels for the during of their period of employment.

Is this another area where the expert panel needs to be informed?

(No response.)

MR. POWERS: Well, fair enough. The question posed and maybe not answered.

Let's move on with the presentation. I guess Mr. Hallbert, are you -- no, I'm sorry. Jay, you're next on the list.

MR. PERSENSKY: Yes, I'm next on the list. I'm going to jump in between Erasmia and Bruce even though --

MR. POWERS: Not to diminish the

importance of your presentation.

MR. PERSENSKY: Just to bring in this human factors element, I'll try to be as brief as possible.

MR. POWERS: Let me say that I did find the slide that showed the coupling between HRA and human factors to be illuminating useful, a point that bears repeating.

MR. PERSENSKY: Well, you're going to have an opportunity to see it again.

(Laughter.)

MR. PERSENSKY: The role as I see it of the human factors research at the NRC is really to provide the regulators -- NRR for the power plants, NMSS for materials, and also now the NSIR -- and their staff with the tools necessary to do their licensing and monitoring tasks. Those tools should be developed from the best available technical bases. With that, there is also sort of an element of maintaining competence with that research to do just that.

MR. WALLIS: Do they know what tools they need?

MR. PERSENSKY: They have an idea of what tools they need because they send us users needs.

MR. WALLIS: Are they specific enough to

tell you what you need to do?

MR. PERSENSKY: In those cases, yes.

The ultimate goal of course is to ensure that nuclear facility personnel have the tools, the knowledge, the information, the capabilities, the work processes, the work environment, both physical and organizational to safely and efficiently perform their tasks. That's generally what we try to achieve.

In your packet I believe you've got a copy of SECY-01-0196, which was the last iteration of what might be called the human performance or human factors plan. That particular SECY said that we were going to in fact sunset the development of a human factors plan or human performance plan as an independent document. Further, that those activities that might come through the human performance program would in fact be incorporated either in the HRA plan or the Digital I&C plan in the future.

MR. POWERS: Now I saw no one crying over the demise of that document.

(Laughter.)

MR. PERSENSKY: And that document also presented where we were at that time.

The next slide is duplicate of the one you saw in Erasmia's presentation. And again, it's just to remind you that there is an interaction, an ongoing interplay between the HRA disciplines here and the human factors disciplines.

We're trying to work more closely both in terms of providing the information and the data so that we can enhance the HRA models, indicating where there might be some problems where we need something but that HRA/PRA isn't able to provide at this time.

One the other hand, they provide us, in doing some of the work that we do, areas that we should be focusing on, the needs that they have for more data, and as well as an opportunity to provide prioritization for the work they do.

This is the relationship between these two groups. It doesn't say that we don't do things on the other side as well, but in fact we do develop things, the tools that they need. There are tools for the HRA but there are also tools for the regulators.

MR. WALLIS: It would be reassuring if you had things coming in and going out.

MR. PERSENSKY: But again, that's how we interact.

We can jump into the next slide, which gives you the listing.

MR. APOSTOLAKIS: Which branch of the

Office of Research is the human factors?

MR. PERSENSKY: It's in the Regulatory Effectiveness and Human Factors Branch. John is our branch chief.

MR. FLACA: I'm am the branch chief of that branch.

MR. PERSENSKY: We are a small team within that branch.

As with Erasmia's slide, you'll see that we do have a listing that's reminiscent along here of the one slide from Scott's presentation, essentially the functions and along the top, the types of applications that you're interested in.

You can see from this that we, again, have a number of activities that are going on. We'll go through some of them.

MR. POWERS: I'd sure like to know what the status is on fatigue.

MR. PERSENSKY: I'll delve right into that then.

NRR has been tasked with developing a rule. One of the reasons for that tasking is that there was a PRM petition for rule-making, as well as we got a couple of letters from some Congressmen. We prepared SECY-01-0113 last year to the Commission that included in it a rule-making plan, and we're in the process of developing that rule-making accords with that plan.

We have almost monthly stakeholder meetings with NEC, industry representatives as well as UCS, and the petitioner are particularly involved.

MR. APOSTOLAKIS: So what you are trying to do here is develop guidance that prevents fatigue of the operator?

MR. PERSENSKY: We're hopefully developing a rule that would allow the utilities to develop fatigue management programs, which would reduce the probability that a fatigued operator -- or fatigued personnel. It doesn't have to be just operators -- would be operating or doing a maintenance task.

The agency currently has a policy statement that was prepared in 1982. And one of the reasons I'm involved with this is I have the unfortunate history of having been the person that developed that policy.

(Laughter.)

MR. PERSENSKY: It allows certain working hours. What we've learned through the years is that working hours is not the only aspect of fatigue.

That's why I mentioned fatigue management programs because over the years, especially in the Department of Transportation, they have been developing new techniques to account for fatigue and way of trying to reduce the effects of fatigue. We're working with the industry to come up some guidance.

The draft rule is due back to the EDO in July of 2003. We're starting the regulatory analysis aspects of that, which is where we need some of the risk information. And as I said, we've been working with stakeholders to come up with some options in this rule-making activity.

You will of course have an opportunity, either at the draft rule stage or the final rule stage, to review that, that work.

MR. LEITCH: In addition to working hours, would this also include considerations of circadian factors?

MR. PERSENSKY: The primary factors that drive fatigue are circadian factors, length of shift, age has a consideration, and the kind of work they're doing. But, there are a number of factors that go into it.

That's why we're trying to do it through this fatigue management aspect, where we may have a rule that addresses hours of work, but there would be guidance, industry guidance of how to train people both for acknowledging and recognizing the effects of fatigue as well as to train others to observe under the behavioral observation program to see if one of their colleagues is exhibiting some aspects.

We've also looked at -- there are some techniques out there. There's some hardware, where you can measure fatigue or keep people awake. We've done some analysis of that. We're not necessarily proposing anything in that area.

There are some algorithms that have been developed, particularly in the transportation industry as to -- you use that algorithm and include the time of day, length of shift, how long they've been working over a period of time, that that could give some indication. We're looking at that as some possibilities.

But right now, the rule is not being driven by, again, that part of our technical bases work that we've been doing.

MR. LEITCH: A lot of plants are going away from eight-hour shifts to ten or twelve-hour shifts. Have you looked at that?

MR. PERSENSKY: The best we've gotten, the best count on that is around 50 percent are at twelve-hour shifts for operators, and either ten or twelve-hour shifts for some of the other people. We have done some work previously that has actually said twelve-hour shifts, if done properly, they didn't reduce operator performance.

One of the big issues of course is there's normal operations and then there's outages. And during outages, there's much more use of overtime and going to the limits that are set currently in their technical specifications. In order to achieve the kind of outage periods, they need those hours.

So, we're trying to come up with - again, we're working with the stakeholders and coming up with some methods that we think will be acceptable.

MR. POWERS: If you were totally successful in developing this algorithm that says okay, here are the fatigue effects, as a function of all these parameters that you suggested might affect things: time, age, etcetera --

MR. PERSENSKY: Right.

MR. POWERS: -- and you feed that information to the human reliability analysis folks, wouldn't that drive them to time dependent PRA?

MR. PERSENSKY: I don't know that I know the answer to that. I don't think because it's not all based on time --

MR. POWERS: It seems to me that if indeed fatigue has the consequence of increasing the likelihood of error in the course of a day that you wouldn't want to just -- because it's collective. I mean if one guy on his shift is becoming more error prone, everyone on his shift is becoming more error prone because the shift all begins and starts at the same time.

MR. PERSENSKY: Well, again, during outages that might be more of the case. But during normal operations, it may not necessarily be the case where everybody is staying. There's usually a replacement for someone that's ill or calls in. So from that standpoint, there is some difference between those time periods.

But I'd prefer to turn the HRA question over to our HRA experts.

MR. SIU: Actually, interestingly enough, one of the discussion items in the elicitation process we talked about for PTS, we did talk about things like the time of day. But in the end, you are where you are when the event hits, so you don't necessarily have to track it.

I mean we're not being asked for a time

dependent result for, let's see, the vessel failure frequency. If you want to know how some notion looks at the annual average frequency, of course, figuring into that average is how often you're in a condition that might promote error in generation.

Again, this gets to back the very simple minded representation of the ATHENA process. If time of day were the only factor that you are concerned about, you look at how likely it is that you're in the window and then what the conditional probability of failure given that you're in that window.

As Jay pointed out, of course, if you're starting to look at interactions across the whole plant and all the operating personnel, that can get pretty hairy. But for the control room, at least that's conceptually how we could address that.

I don't know, I guess in the short answer, that that in itself would call for time dependent PRA. It's more, do you need a time dependent answer to address the concern you've got.

MR. ROSEN: One of the important factors I think might be -- since we think that crew performance is very important and not individual performance in the event of an accident or quickly moving scenario -- one of the important factors is, is the crew actually the crew that trained together? What percentage of crews are crews that are actually relieved where one or more of the members are not part of that crew, or have been socialized or trained with that crew?

This could be important. Is this something that you're looking at?

MR. PERSENSKY: Not necessarily with regard to this particular effort.

MR. ROSEN: I suggest you think about that as part of what you do.

As long as I'm interrupting the train of thought, when you get the risk-inform CAP, I'd like to hear about that although you didn't underline it. I'm not sure what underlining means in this chart. I guess it means you're not going to talk about it.

MR. PERSENSKY: No. What it means is that I attempted, but failed because of my lack of knowledge of Microsoft to make this a linked presentation where I could just click on that and it would take us to the appropriate slides.

(Laughter.)

MR. PERSENSKY: It works fine on my computer. And if you'd all like to go up to my office -

(Laughter.)

MR. PERSENSKY: Unfortunately, when you put it on to an "A" disk, it loses all those links. I tried to actually come up with a way of fixing that last night except my laptop died at home so I couldn't do that. So, the only underlining was that it was linked.

MR. ROSEN: So you're going to tell me about risk-inform CAP at some point?

MR. PERSENSKY: Yes, we will get into that.

MR. LEITCH: Just further on Dr. Rosen's point, I've been aware of a couple situations where not only didn't the crews train together, which I think is an important factor, but in one case there was a situation where the operators were operating on an eight-hour shift and the operator of supervision was operating on a twelve-hour shift.

So by definition, they couldn't have trained together because for the first eight hours this guy was there supervising, and for the last four there was another supervisor. I mean there's just a lot of this around the industry that just adds to the complexity of the situation.

MR. PERSENSKY: It is a complex situation, and we do not regulate currently in terms of the number of shifts or the way they rotate. We do regulate the number of licensed operators that are required on each shift, depending on the mode that the plant is on.

But, we don't tell them that they have to rotate together. We don't tell them they have to have six shifts. It sort of works out that five or six shifts works out to be a good way of running it unless you have 12 hours then you'd go down to four rotations.

Each plant does have its preferred way of doing it, and at this point we don't regulate with regard to that.

MR. ROSEN: But if you found a way of doing it that had negative risk implications, I assume you would regulate it, wouldn't you?

MR. PERSENSKY: If we could determine the actual effects from a risk perspective. Personally, I don't think that risk models at this point are mature enough to be able to do that. I may be wrong.

MR. ROSEN: Is that what we're trying to do, to find out what is it about human performance that's positive and negative, and reinforce the positive, and do things to not let them get into negative conditions. That seems to be the whole objective of this thing.

MR. PERSENSKY: That is the general objective.

MR. ROSEN: So I would encourage you to be thinking about training and crew performance in that light. There are some things one can do in a power plant in terms of staffing the control room that are not good from a risk standpoint.

MR. LEITCH: This particular situation, I just found out is not good from a risk standpoint and I had it changed. But what I'm saying is it had been going on for quite some time. Intuitively, it doesn't seem to make sense that for some portion of the shift you're reporting to one group and --

MR. ROSEN: And what drives that is absenteeism. I mean plants don't set up to have a lot of that kind of thing happen, but it happens in fact, especially in plants with very experienced crews that need to have a lot of time at the plant, which means their older, they have more vacation -- these programs add vacation for people as they get 10, 20, 30 years of employment. That means that the guys is not necessarily sick. He's just taking his vacation. And when does he take his vacation? When the plant is not in the outage. In the outage, they try to get everybody to come to work.

So during normal operation, you're going to find many, many crews with people who are relieving crews that are not self-relieving, where you don't have enough people to fill in on the crew, with people who have trained with that crew.

So you're going to have lots of circumstances in which the crews haven't trained together even though we all know it's best that they do. In fact, they do train together. Our simulator tests are based on crews that are training together. So in that sense, they again can confound the analysis. If you use the data from those tests that confounds, it's not going to be as good as that in the real world because of this phenomenon described.

These are human factors considerations. I'm just mentioning them because I think they're important.

MR. PERSENSKY: Thank you. And in fact, it does encourage certain things. But again, it's more of an encouragement rather than a direction.

SRP Chapter 18, again, this is a tool. This is a real tool that the people in NRR use. This is a human factors chapter. It's based on the document that we prepared, NUREG 0711, which is the human engineering program review model. It addresses how we should do our reviews of changes to our plans, to new power plans, our control rooms.

We've done a number of projects related to bringing together enough information to go forward with the revision to that SRP. Again, that will be subject to an ACRS briefing.

MR. APOSTOLAKIS: Does it get into organizational issues?

MR. PERSENSKY: It does not. Chapter 18 does not get into organizational issues. Chapter 13 does have some element of organizational issues. But, it's not in Chapter 18.

Chapter 18 focuses primarily on interface. It's a process kind of document. It also has some aspects of procedures, training, and all that in to how you would do an entire human factors program at a utility.

I mentioned earlier the staffing work. The project here, again, this is based on user need that relates both to advanced reactors as well as to current reactors in that some reactors in their changes -- you know what I'm saying? If we completely change out our control room as completely a digital and much more automation involved, there would be some opportunity to perhaps reduce the staff at a conventional reactor.

We're trying to develop a tool that would be used by the licensees that is based on what is called Path Network Modeling, which is a type of human behavioral modeling used extensively in the military. Also, NASA uses similar models.

We have done some testing of this type of modeling in the past in terms of trying to say, how good is it, by doing experiments where we have a shadow study, where you model and see how well you think the operators would perform. Then, actually collect data at a simulator to see how well the operators do perform given the various situations that could addressed to try to verify or validate that modeling technique.

At this point, we're looking at trying to develop this as a tool for the review of staffing proposals that come in from the utilities.

MR. LEITCH: I guess I'm trying to differentiate between new reactors and current reactors. Are there any plants where the licensees are seriously proposing changing their control rooms because of instrumentation changes?

MR. PERSENSKY: We have no applications. Again, because I'm familiar -- I've been working with EPRI on their development of design guidance for hybrid control rooms -- that issue has come up a couple of times at those meetings as possibilities.

MR. LEITCH: You also said something about additional automation, if I understood you correctly.

MR. PERSENSKY: That's right. Those are things that are being considered by various utilities at this point.

MR. LEITCH: being considered for current plants?

MR. PERSENSKY: Current plants. There is at least one plant that intends to shut down and completely replace their control room at one time. as opposed to doing the piecemeal type of changes that have been mentioned.

MR. POWERS: When you bring up the issue of automation, there's also the issue of non-automation. And with existing reactors, it seems to arise in front of the ACRS episodically, but maybe a cycle of every three years, where the issue comes up: should we automate some function because there's insufficient time for manual action?

The staff has at various times attempted to approve I guess it's a regulatory guide in that regard, and the committee has resisted it because the underlying data is proprietary. Is there something being done to address that situation?

MR. PERSENSKY: Actually, there's a NUREG that has come out -- unfortunately, I can't remember the number off hand, but I do have it here -- that attempted to come up with a different method whereby you would use risk information to categorize the risk level of a particular operator action. Based on that, they would determine the level of human factors review.

Again, that will be part of the Chapter 18 revision. You'll have an opportunity to see that in more detail when that comes for review. But, we are looking at that as a replacement for ANS 58.8.

On the reactor oversight process, the ROP, we did a study -- actually, INEEL did the study for us -- on looking at whether or not the reactor oversight process adequately address human performance or what kinds of things may not be caught given the reactor oversight process.

A major recommendation that came out of that particular piece of work was that it appeared that a number of the corrective action programs were not keeping up to date with -- they weren't able to implement the fixes rapidly enough or prioritizing the kinds of fixes based on risk. So, we were seeing repeat kinds of incidents.

So, we recommended to NRR that they look at the current corrective action program inspection module, which essentially asked the review or inspector to use risk as one of the aspects of looking at what they should be reviewing. But, it doesn't give them very good guidance to what that mean.

We proposed to NRR that one of the things we'd do is to provide better guidance on how to do that, that risk of the backlog in the corrective action program. We have not heard back from NRR on that, but that's one of our recommendations.

MR. ROSEN: I think that's a very valuable step. Although, I've seen some very good corrective action programs in use in utilities.

There is still that weakness that they don't prioritize very well based on risk. The priorities are more historical in context. Maybe the highest priority things are things that are reported on LARs.

There are different protocols that are not risk based for prioritizing work in the plant. I think that's a fundamental flaw. I've encouraged some utilities to do better, to do what you're suggesting or at least consider risk as one of the primary things that you think about when you prioritize corrective action.

MR. POWERS: How could you do that if you don't have a fire PRA?

MR. ROSEN: Well, fire is not the only risk. But I think in the cases where you have a fire risk and don't have a PRA, it's a problem.

MR. POWERS: Sure. But I've seen based on my episodic trips to plants in examinations of corrective action programs, if I'm going to guess what is the longest, the corrective action with the longest lifetime on list, it'll always be something connected with the fire protection system.

MR. APOSTOLAKIS: I was reading the root-cause analysis that was done for the Davis-Besse incident, and in several places there are sentences like "plant was restarted without taking corrective action for identified problems" and "the management ineffectively implemented processes".

Are you trying to help the corrective action program from that point of view? I mean what do you do if they know about the problems and just don't do it?

MR. PERSENSKY: That's part of the inspection process. We're trying to develop a way to at least identify what they should be doing or what are inspectors should looking for.

At this point, and I'd have to turn it over to NRR for a regulatory perspective as to what decisions they'd make. Those are regulatory decisions that they'd have to make.

MR. APOSTOLAKIS: Is there any attempt to develop performance indictors or good corrective action programs verses a bad one?

Another thing that was missing evidently was doing hazard analysis. That seems to me to be something that one can look at the work processes and identify. Incorrect implementation of a program is not an issue of a work process. It's something else. So, I wonder whether it would be a good idea to try to develop some indicators that will alert the inspectors to the fact that something is not being implemented right?

As you know, the reactor oversight process, a good piece of it is performance indicators. Well, these performance indicators have nothing to do with human performance.

MR. PERSENSKY: That's correct.

MR. APOSTOLAKIS: But the question is should we be trying to develop performance indicators for human performance, not necessarily of the same kind where they have frequencies or events, but maybe of some other kind but, still performance indicators. Or, should that question be addressed to NRR? I don't know.

(Laughter.)

MR. POWERS: Tell him the answer is NRR. Let's stay on human factors here.

MR. PERSENSKY: We do have some STPs. There's an STP on licensing for instance, and there have been some attempts in developing further STPs in the human performance area.

Most of those things that you would pick up in the human factor area come out of inspections, not out of the PIs. The assumptions were that the PIs would be something that would -- human performance would show up in the PIs. That's why they call it a cross-cutting issue.

So, we have not yet attempted to do a human performance PI. Back in the early 90s, we took some shots at it.

MR. ROSEN: I think they should. A lot of plants have much better human performance data than they used to have.

MS. LOIS: I'm talking about late 80s, early 90s.

MR. LEITCH: The problem is that there's no uniform standard as to how the plants collect and analyze that data. I mean every plant has its own system of doing things, some of which are very effective. But when you compare plant A with plant B, it's very difficult to perform that kind of comparison.

MR. APOSTOLAKIS: But I think one of the things that should be done, probably by your group, is to look at the inspection, the ROP, and take the root-cause analysis of Davis-Besse and other analyses, and every time they identify a problem, ask yourself: which part of ROP would actually catch this? Some of them are easier to catch than others.

MR. PERSENSKY: In a way, we did. What we did in doing this project was we went back to ASP reports -- or ASP plants that were high-risk plants, and looked at whatever archival data that we could then compared it to the ROP process. But of course, most of that data came from pre-ROP events. I think to follow on with some more recent situations like

Davis-Besse and Indian Point might continue to be an exercise.

MR. APOSTOLAKIS: I think that would be a very good exercise.

MR. PERSENSKY: But that exactly was the process we used.

I mentioned some of the other work as far as the inspection manual for the materials and waste area. Erasmia mentioned fitness for duty. Fitness for duty as you know is undergoing a rule change. They're talking about including fatigue and decommission of plants in the drug and alcohol portion of fitness for duty. In fact, fatigue is going to be in part 26 of the rulemaking. There won't be a separate rule for fatigue. It's probably going to be in part 26.

Just a couple things on what we consider to be infrastructure of the development of the needs to support the other work. The Halden Reactor Project, which some of you are familiar with, is one of the few places that we have access to simulators for research projects. We're been using the Halden project, that project in Norway.

MR. POWERS: I've got to ask my questions.

MR. PERSENSKY: I knew you would.

MR. POWERS: I still have to understand how a Norwegian reactor operated by a Finnish has an yield results that have any applicability to American reactors operated by American crews.

MR. PERSENSKY: To start off with one is to correct some information. One, it is a simulator of a Finnish reactor and we use the crews from that plant. It's from Loviisa, so they're Finnish operators operating a plant in Norway. They happen to be located in Norway, but they're inside an enclosed building. It really doesn't matter. And, they're used to the weather.

(Laughter.)

MR. PERSENSKY: As far as trying to give just a briefing, we have looked very closely at what goes on. We have looked at their training programs, we have looked at their procedures, and we've compared it to the kinds of things that go on in the US. But the bottom line is it's something that's available to us. We don't have a research simulator here in the US. That is something that we can modify as we can with the Halden reactor.

MR. ROSEN: Should you have a simulator for research here in the US?

MR. PERSENSKY: I think from the

standpoint of cost, it would be fairly hard to justify.

MR. ROSEN: Don't get into the cost.

MR. PERSENSKY: It would be very useful to have our own research facility. We've addressed this. Actually, there was a DOE meeting earlier this year, I guess in May, where we talked about it in terms of developing a research simulator for advanced reactors.

MR. ROSEN: I was thinking of a multi-capable simulator that you could configure.

MR. PERSENSKY: That's exactly what the Halden simulator is. In fact, we talked about that in the past, but they now can configure it to be used as a PWR or a --

MR. ROSEN: It's basically just a computer, right?

MR. PERSENSKY: It's a computer with some workstations.

MR. ROSEN: Right. And the more you get towards an N4 type control room, where the operator sits in front of a computer screen, the easier it is to change the program and then you're in a different plan.

MR. PERSENSKY: Right.

MR. ROSEN: It would seem to me that one

ought to be thinking about that sort of thing and not saying we have to go to Norway and use Finnish crews because we don't have that in the US. What we have in the US is what we need, and if we need it then we ought to be thinking about it.

MR. POWERS: I think it's an excellent question. It's exactly what this subcommittee ought to be pursuing, what would be very desirable to have. It's what the people like John Flaca get paid the big bucks for to decide what they can actually afford do. And the Commission gets big bucks to decide where the money ought to come from. But we ought to be deciding what would be desirable.

MR. ROSEN: We ought to be at least discussing it.

MR. APOSTOLAKIS: Yes, as long as they promise not to fly over the Finnish crews.

MR. POWERS: Yes, don't bring the Finnish crews here.

(Laughter.)

MR. PERSENSKY: They're interesting people. Bruce has had a lot of opportunities since Bruce actually worked in Halden for several years. He did excellent PRA work as a matter of fact.

MR. POWERS: And the Swedes make excellent jokes about them too.

MR. PERSENSKY: But currently, it's part of our infrastructure. A big part of it is the fact that they have a facility that we can use that is reconfigurable. And we're moving towards making better use of that data for HRA, not just human factors projects.

MR. POWERS: That's really the substantive issue. That, you've collected all these data from the Halden project, now what do we do with it?

MR. PERSENSKY: We have used it in the past for the development of the guidance that is going to be in the SRP.

MR. POWERS: The question often comes down to is that the source of the three-foot telephone cable and --

(Laughter.)

MR. PERSENSKY: There never was a three-foot telephone cable. That was a miscommunication. We have gone back and looked at all versions, draft version of those 700 and there was never one that included a three-foot telephone cable as a guidance document.

MR. APOSTOLAKIS: Humphrey Bogart never said play the games. You're suffering from the same thing.

(Laughter.)

MR. PERSENSKY: Actually, now the question is whether or not we're going to allow wireless.

MR. POWERS: In seriousness, Steve has raised the question: should we have our own research reactor? I mean has this Halden thing proven so useful that in fact we should have our own? The question is, indeed, are others' data proving to be very useful?

MR. PERSENSKY: We have made use of the data. We intend to make more use of it, especially in the HRA area. That doesn't necessarily negate the question. Again, part of it is just like everything else. It's a cost/benefit issue.

MR. POWERS: Yes, but other people in higher pay grades than ours get to make the financial decisions. We ought to be making the technical decisions.

MR. PERSENSKY: Well, we have --

MR. POWERS: . . . go by the committee and sell us three times over just based on his monthly wage, right?

MR. FLACA: Well, the question comes down to what is the benefit of going in that direction above and beyond what we can get from the reactor. That's the bottom line.

We do have simulators at TTC. And to move ahead and look at advanced reactors, I mean it has really established the capability to be able to ask questions. Whether or not we're asking all the right questions -- we might be, but how do we know for sure -- there's still the uncertainty that surrounds that aspect. And, the question is how does it indicate, or what kind of indication, or how much can we gain from something that we own verses something that we observe and move into collaborations with other organizations?

You're right, it has to be thought out. We need a basis for going in that direction. That's up to the committees. We need insights in those kinds of issues. It's very helpful to us in making those decisions. I think that's why we're here.

MR. POWERS: Yes, I mean I just like the idea that there'd be some vision or -- I appreciate Steve bringing the question up.

MR. FLACA: Yes, sure.

MR. POWERS: The issue that most perturbs me about the HRA and human factors areas is this vision of what we really ought to be as opposed to what -- all these day to day activities that we're carrying out right now, what do we really want to be in future in this area? That, I don't have well articulated. I mean I don't see the vision right now, and it's going to come up this afternoon when we discuss tools.

MR. ROSEN: Dana, it will come up because Peter Ford is asking it. In the context of ACRS' review of the advanced reactor research program, he has asked the question: where do we want to be in 15 years? And I think in the human factors area, we need a whole new set of questions.

It's helpful for me to go through this dialogue with you and the rest of the committee because we need to answer that question. You and I, Dana, have to write that section -- you and I and several others.

MR. PERSENSKY: Just to finish up my part, I just want to touch on something because it also addresses the issue of words we can't say like safety culture.

Under international activities, one of the things that we did agree with the Commission is that we would be able to follow what's going on in other places. To that end, we have Dr. Shurston Dahlgren from IAEA, who is one of the people that does the safety culture reviews for the IAEA, who will be giving a seminar here on September 23rd.

MR. APOSTOLAKIS: Is that the ASCOT methodology?

MR. PERSENSKY: More than that. It's gone beyond ASCOT.

But, she's coming here and will be giving a seminar on --

MR. APOSTOLAKIS: Who is this person?

MR. PERSENSKY: Shurston Dahlgren.

MR. APOSTOLAKIS: Oh, yes. I know here.

MR. PERSENSKY: At 10:30 and --

MR. APOSTOLAKIS: Which day?

MR. PERSENSKY: September 23rd. It's a Monday. September 23rd at 10:30 in T-10-A1 of this building. It went out as a network announcement. Do you guys get the network announcements?

MR. APOSTOLAKIS: Oh, yes.

MR. PERSENSKY: It said "seminar on safety cultures".

So again, that's part of what we're doing in keeping abreast of what's going on. Since we're going to have this afternoon to get into more detail on some of these things, I'd like to turn it over to

Bruce because they need to leave this afternoon to go back to Idaho.

Bruce is going to talk about a project or a couple of projects really of how they have taken Halden data and are trying to apply it into the HRA. Bruce was at Halden at the time the work was being done.

So, Bruce Hallbert.

MR. HALLBERT: Thanks. Can I borrow your microphone?

MR. PERSENSKY: Sure, if I can get it off.

MR. APOSTOLAKIS: Human and intelligent systems. There is a clear distinction between humans and intelligence.

MR. HALLBERT: It's not meant to be exclusive, George.

MR. APOSTOLAKIS: How do I know?

MR. HALLBERT: Good morning. I'm Bruce Hallbert and I'm pleased to be invited to speak here.

As Erasmia and Jay have mentioned in their discussions, we're doing work with the Nuclear Regulatory Commission in the area of human reliability analysis data. I'm going to talk this morning about using simulators in human factors research with the subtopic of linking this human factors research with human reliability.

There certainly are a variety of sources of information that can be used to form human reliability analyses. It's the hypothesis of this discussion here that simulators are one of those viable sources.

Next slide please. As Jay mentioned, most of the work that will be presented in this discussion was work that was conducted while I was in the Halden Reactor Project although some of the sources that are referenced here were also generated by the INEEL previously.

So the purpose of the work being presented today is to discuss the study of human performance in which data are present to inform HRA activities. I'll discuss more about that study in the following slides. But the intent in doing so is to illustrate, for example, some of the relationships between human factors research and HRA to show that they are complimentary and can not only co-exist, but be very fruitful in their interactions.

Next slide please. I'll start the discussion today by discussing some of the potential areas in which simulators can support, or where simulator-based research or activities can support human reliability analysis. Then I'll move to an overview and a background of a particular simulator-based research project.

This project was sponsored by the Nuclear Regulatory Commission to evaluate the issue of main control room staffing for advanced reactors. I'll talk specifically about what was the issue under consideration and what we mean by, specifically, what kinds of advanced reactors.

I'll provide a background to that. I'll talk about how we did it. I'll talk about the underlying science and assumptions that were important in guiding the way that we set up the experiments, which data was collection. I'll give you some examples of how those studies were conducted, including pictures, then talk about the results.

The results from the study that I will be presenting will be relevant for the issue of staffing of advanced reactors, but I hope to use it to illustrate the convergence of that particular research topic with the general topic of human reliability analysis.

From there, I'll move into what we're calling an embedded study, which is a preliminary exploration of performance shaping factors and performance, specifically main control room operator performance with the notion that there's linkage here between studies of performance shaping factors and operator performance and HRA. Then, I'll summarize the results.

Hopefully then, where we want to go with this is to have sort of an open discussion on the potential of these kinds of things in supporting HRA.

Next slide please. It's our position here that simulator studies and simulator-based activities, whether they're studies per say or not, can provide useful data for HRA.

By that we mean, for example, you can carry out research embedded within other activities in which you can explore the relationships between performance shaping factors, which are an important element of human reliability analysis methods and performance and hopefully also by extension to consider situations of operator error.

MR. POWERS: Mr. Rosen has raised the issue that seems to me to strike at the heart of this hypothesis that you've put up here.

MR. HALLBERT: Yes.

MR. POWERS: That, the thing that upsets the performance of a crew the most is when we have an interloper in here.

MR. HALLBERT: A what?

MR. POWERS: An interloper, someone who has not trained with this crew, who has not socialized with this crew. I mean it's not like we got a Finnish operator and stuck him in here. But, he is different.

It seems to me that until you can address Mr. Rosen's question, this stands subject to some substantial debate.

MR. HALLBERT: Okay, I'll be happy to entertain that debate as well too. I intend to address the issue of making conditions representative for making inferences that are applicable to US plants from these kinds of studies.

The specific issue of the interloper --

MR. ROSEN: Well, I think Dana is maybe exaggerating the importance of it.

MR. HALLBERT: He'd never do that.

(Laughter.)

MR. ROSEN: I think it's important. But, where a qualified SRO, for example, relieves someone from the crew who is on vacation, and he is from a different crew, perhaps on his weekend so there is a fatigue consideration because he comes in at a time where he's supposed to be resting.

He has not trained with this crew and different communication protocols were established perhaps -- not fundamentally. He's still taking part in the three part communication and that sort of thing, but he may not be in his normal role since he may be operating as unit supervisor. And, in the crew that he's actually in, he's just a SRO or vice-versa.

So, there are also different issues of how people communicate, who's in charge here, what do you expect me to do, what do I do --

MR. POWERS: But that unusual circumstance is never going to be reflected in the data they get.

MR. ROSEN: Right, it's not. And that's the question I pose. Is it, and how would one address it?

It's a fairly normal circumstance. I would guess that in plant on average -- now this is just a guess -- but perhaps 20 percent to a third of the time.

MR. POWERS: My calculation said it could be as high as a third.

MR. ROSEN: High as a third. So 20 percent to a third of the time, you'll find crews operating with one or more members who are not part of the standard crew.

MR. HALLBERT: I think the condition that you are describing can be studied through simulators. I think simulators would be a very logical way of evaluating that particular issue through the collection of data.

I'll say also that I don't have any data here to present today, but my own personal observations from having conducted a number of different research projects like this would be consistent with the issue you raise here. That, in fact, team performance is critical and the factors that contribute to that, if they come out of alignment with regard to leadership with regard to communications factors and the normal division of labor and aspect like that, can influence performance and have influenced performance.

MR. POWERS: My next question is having identified one potential flaw in the use of simulator data, what are all the other flaws?

MR. HALLBERT: All the other flaws of using simulator data?

MR. APOSTOLAKIS: Why is it a problem?

MR. HALLBERT: I don't know.

MR. POWERS: But see, we're calling into question all the simulator data --

MR. APOSTOLAKIS: Or that exists.

MR. POWERS: That exist, that can be generated. If I can do with one question raised by member of the subcommittee here spontaneously, are there lots of other things?

MR. BONACA: I have other things. The question I have is -- I mean this is being done for foreign plants. But do you have crews just as they are in the US? The question is do they have written procedures as we have in the US, which are different from procedures in other countries?

Those are really questions that I think will really affect the performance.

MR. HALLBERT: Yes, let me address those head on. I was going to address them in some slides that are going to come, but I'll take them right now.

It was, of course, a concern for us in designing this particular study but other studies as well to make the results generally valid, externally valid to the user group. In this case, the Nuclear Regulatory Commission in the US.

What we did to address some of those concerns was that we traveled to the plant in Finland that volunteered to participate with us in the study. We had the NRC along on that trip, and we evaluated a number of things.

We looked at their training program and found it to be generally comparable to IMPO standard accredited types of training programs for training licensed reactor operators and other control room personnel. So, we looked at that and satisfied ourselves that they were following a process similar to what US plants follow for training their personnel in the control room.

We looked at how the division of labor was accomplished in the main control room because this was a study of main control room staffing. Again, we were satisfied that the division of labor fell into the same major categorizes as in the US plants and very closely, parallel to division of labor of control room personnel.

MR. BONACA: And they have symptom-oriented procedures?

MR. HALLBERT: Yes, and I'll come that in a second. I'll finish with the staffing though.

They have a control room supervisor, who may also be the shift supervisor. They have shift technical advisor, who is also a degreed engineer who has also got training in reactor operations and license. They have a balance-of-plant operator and their philosophy for control room operation is similar to the philosophy of control room operation at the US plants.

MR. POWERS: You say that it's similar?

MR. HALLBERT: Yes.

MR. POWERS: That means that it's not identical. How do I judge similarity? I mean how close is close.

MR. HALLBERT: I would say in similar and all relevant aspects that would contribute to the findings from operator performance in generalizations to the US situation here. In other words, they were so similar that we couldn't really detect any meaningful differences.

There are some differences in the plant design, of course, so we couldn't say that the function allocation or all the responsibilities for this reactor operator at the Finnish plant would be the exact same as those for the US plant operator because there are these plant design differences.

MR. ROSEN: There are plan design differences in the US as well.

MR. HALLBERT: That's true.

MR. PERSENSKY: As well as control room operating philosophy. I mean we just talked about it before that in some plant they train together and in others they don't. They rotate together and in others they don't.

So, there are differences with in the US. I don't think that the differences that we observed at Loviisa were that much different than what you would see within plants here.

MR. BONACA: What about cultural --

MR. HALLBERT: For these intense purposes, I think --

MR. POWERS: What's causing the question is -- you're going to collect simulator data and you're going to say, from this I'm going to make judgments about normal operations. We've identified one potential flaw in that data. I don't know that it's a flaw, but it's a potential flaw. And now we come to this cultural flaw and say that it's similar, but we know we have a vast amount of differences. So, it's only similar to some subset of US reactors. It lacks generality.

These are the kinds of questions the research program has got to be generating concerning its experimental methods. I'm questioning whether we've done an adequate job here. I'm questioning their methods.

MR. HALLBERT: We certainly had the same concerns at the outset of whether we could the study for the NRC. That's why we had the NRC along with us at these meetings.

I guess maybe what I should say is that where we ended up on the issue of main control room staffing and division of labor and responsibilities is we found them to be equivalent from everything that we had to compare them by.

MR. BONACA: One thing that is known about Loviisa is they really have an outstanding history of operations, technical management, and extremely involved crews. I'm not sure you're going to reproduce that kind of quality. All I can say is that from what I understand is it's the kind of performance on their part.

MR. HALLBERT: One of the reasons why we selected them was that they had set world records for availability and performance, and also because they were very advanced within the European countries for their use in PRA and incorporating it into operations and procedures.

MR. KRESS: It seems to me like your studies are asking the question: Is this something that would be a useful approach? It may not be definitive in the detail of quantifying it, but if you made the judgment that this was an approach that is useful then to address these questions of differences in culture and differences in plants, it seems to me like you would need to go to actual US plant simulators with US operators and do this same sort of study on a plant specific basis across the country.

Is that something that's part of the thinking if this proves to be a viable approach?

MR. HALLBERT: I think that's a good idea. There was previous research that was done and the author of the work was Ed Marshall. He considered all the factors that could contribute to confounding of results from simulator-based studies or experimental research at this time.

So, there had been some thought previous given to that. We used that work that was done -- and I would apposite that for future work of this kind that some kind of list like that or methodology for consideration of confounding factors needs to be taken into account.

MR. BONACA: That goes to the heart of my question too of why we haven't talked yet about the symptom-oriented procedures because that included all these elements. That includes all the elements of observation of simulators, tailoring the procedures for specific situations, and in fact, testing to verify that those kinds of estimations and reactions, etcetera, were correct.

The other thing is that procedures went heavily into abnormal conditions and really no design for the situation as you recall from previous observations -- so, there is a lot of valuable information.

I've always felt the pressure because of the timing. I mean every year that goes by that we don't have the information, the vendors are going to lose it because the people in those companies are going away, they're not there anymore. I think having that information would be a tremendous benefit to these activities. I'm not saying that you should just take what is there.

MR. KRESS: Is that information sufficiently complete to form shaping factors and their quantification?

MR. BONACA: Well, I remember for the BWRs there were a number of iterations to the APGs that went year after year. We worked for years doing that kind of stuff. Some of them that were tested weren't acceptable. Therefore, there was a new generation of

APGs. The data placed in the industry around the APGs were the same for the PWRs, were extensive.

MR. PERSENSKY: One of the big problems with any industry data is its availability to the NRC.

MR. BONACA: I understand.

MR. PERSENSKY: Just as Dana brought up earlier, the ANS 58.8 was based on work that was done for EPRI. Because of its proprietary nature, it's not available so it's hard to get at a lot of that data.

The same thing is true with using utility simulators. One, mostly they're busy. And two, they aren't that eager to allow NRC to come and do research.

MR. BONACA: EPRI also generated the scenarios that you have assumed in the back of the procedures, the technical portion. The rest was developed by the ORE groups. Much of the information was in the hands of licensees. And I think they do need to share it.

MR. HALLBERT: I think that's a good suggestion. I think in our first consideration of what are the potential sources, we shouldn't leave stones unturned. We should try to take into account what data is out there. Even if it doesn't suit the purpose that we're looking for right now, it may suit another purpose in the future.

MR. ROSEN: And I think you should be careful about saying that because it's proprietary, they won't give you access to it. All that means is you can't have it in the open literature. Typically that means you can't ascribe the data to a specific plant. But if you want it and went to the right place at EPRI, they might agree to give it to you.

MR. PERSENSKY: In fact, that particular ORE data, we did get access to. But again, there's difficulty in making it available to others and to reference it because of the --

MR. ROSEN: My only concern is --

MR. POWERS: Let me interject here.

MR. ROSEN: -- because it's proprietary, that doesn't mean you can't get the value of it if you approach the problem correctly.

MR. POWERS: The problem is when you try to use it for a regulatory process, you have to give it to the public.

Let me just interject. You have a time limit. You're on slide 4 of 17. I intend to interrogate the committee, which is tough, and we'll go through lunch with no trouble at all but they get to be irascible.

MR. HALLBERT: I'm happy to work to your schedule here.

MR. POWERS: I think that's fine, but I suggest that we go on through this presentation because I to want to interrogate them and then break for lunch.

MR. HALLBERT: Okay, that's fine.

MR. APOSTOLAKIS: One last question. I see here that you're planning to investigate relationships between PSS and so on. One of the major criticisms of the EPRI simulator data was that I believe they tried to come up with numbers, probabilities of human error.

Are you going to do the same? I do like this testing of hypothesis and the relationships. In other words, the structural part -- maybe the simulators will be extremely valuable there. Are you planning to go all the way to the numbers or stop short of that and switch to modeling?

MR. HALLBERT: The numbers that we generated in this study were used for modeling, developing a predicted model and evaluating or at least starting some preliminary thinking on what you could do next with it, but with the notion in mind of trying to better understand the context in which performance shaping factors drive performance and why that is to better inform human reliability analysis. Then also, with the notion or the question of if I have established a relationship between performance shaping factors and performance, what will it take me to establish a relationship between these same things in error?

MR. SIU: I'll take a whack at it also George.

My guess is that we could conceivably generate numbers or a limited number of situations, of course, where the error force in context is strong; therefore, the error probability is high enough that you're going to get observations. There will be other places where we will have to rely on modeling. That's where having these more fundamental relationships between say PSFs and error would be helpful.

MR. APOSTOLAKIS: I like that. I think that's a good idea.

MR. HALLBERT: So if we could jump ahead to the next slide then. Is it my understanding that you want me to finish my talk in five minutes then?

MR. POWERS: Yes.

(Laughter.)

MR. POWERS: No. We want you to take

whatever time you need to finish your talk. The committee is used to working long and late hours and what not. They're tough. I'm not worried about them. I'm worried about the speaker.

MR. HALLBERT: All right. I appreciate your concern.

So, I will then move on to the portion of the presentation and provide some background to the particular setting in which the human factors research was conducted. And that was for a study of control room staffing levels for advanced reactors. That's the title of the NUREG that you see referenced at the bottom of the slide. It's NUREG/IA-0137, published in 2000.

MR. APOSTOLAKIS: What does "IA" mean?

MR. HALLBERT: International agreement.

MR. POWERS: Didn't the committee get copies of this?

MR. APOSTOLAKIS: I don't remember seeing that.

MR. POWERS: I mean I know I got copies.

MR. HALLBERT: The Nuclear Regulatory Commission had received submittals from several advanced reactor plant vendors. These included the AP 600, the GES-BWR, the ABB plant, and the Cando 3 plant.

In these submittals, there was some variability in the proposed changes for control room staffing. That put the issue squarely in the area of 10 CFS 50.54 (m) and changes.

The vendors sited improvements in ease of performance through primarily passive system design and automation as being the primary reasons for requiring a reduced main control room staffing compliment. Some of the pictures in there showed one reactor operator overseeing several plants. Most of them showed a crew, like a modern crew, in a plant control room.

The issue then became one of trying to better understand the performance implications of staffing and advanced plant performance because it wasn't simply a matter of changing a control room, going from a conventional control room to an advanced control room. You were also introducing greater automation and passive system performance.

So, we set out to conduct a study of control room crew performance, recognizing that in order to do so, we would have to establish an advanced and conventional plant benchmarks. And by that, we were concerned very much with the notion of crew staffing and what would be the appropriate references for different staffing compliments as well as thermohydraulic performance and automation.

We developed a range of design basis scenarios, including two involving loss of tools and accident, that were a steam generated tube rupture and an interfacing system where sequence V ISLOCA --

MR. APOSTOLAKIS: Is ISLOCA a design basis?

MR. HALLBERT: It's a sequence V in a PRA.

MR. APOSTOLAKIS: But it's not a design basis.

MR. POWERS: No, it's not.

MR. HALLBERT: Okay.

There was a loss of feed water, a loss of oxide power, and a stem generator overfill. So, we had undercooling as well as overcooling transients representative as well too.

The thermohydraulic performance reference benchmarks, we obtained from previously funded NRC research identified in NUREG Contract Report 4966, which looked at a variety of different transients, overheating and overcooling and LOCAs on BNW, combustion engineering, and Westinghouse plants.

For the staffing configurations, we looked at two different staffing configurations: a normal and a minimum staffing configuration. This whole study was carried out at two different simulator facilities. One was at the Loviisa Nuclear Power Station Training Facility in Loviisa, Finland, and the other was carried out in the Halden Human Machine Laboratory in Halden, Norway, which represented the advanced plant.

Next slide. In the next two slides, I'll go into some of the particular of the study. I think I've talked about these a little bit earlier.

For the phase of this study that was carried out at Loviisa, we looked at the thermalhydraulic performance at the Loviisa Nuclear Power Station to the simulator transients under consideration here. And we recognized, as you might well expected, that there were differences in the plant performance compared with western plants.

Primarily, the Loviisa plant had longer time constants for the overcooling and overheating scenarios than the western plants. They have 16 generators with larger inventories and capacities so they respond a little more slowly to some of these accidents.

What we did was we worked with the

simulator facility staff to modify the simulator from a hydraulic performance to bring it into the range that was more consistent with US plant performance for those same simulator transients. As you might well suspect, that would introduce a confound in the experimental design. So, we then also had to compensate for that by giving the operators additional training prior to participating in the study in Loviisa, Finland, and getting them to a similar level of performance since they would experience otherwise.

The crews in this study operate as crews in the plant. We didn't pull together people from different shifts based upon availability. We designed our study around the availability of crews as crews. We wanted to have actual performing crews.

As I mentioned to you earlier, we also evaluated the training programs and their control room staffing compliments and found them to be equivalent to what we saw in US plants for those features. The thing I didn't have a chance yet to touch upon was the procedures. I'd like to address that now.

In discussions with the Loviisa plant staff, we reviewed their emergency operating procedures. And I hope I represent this correctly, but I believe that they had a previous project or contract with a western vendor, a US vendor, and had undergone the development of symptom-based procedures at their plant. When we came there, the procedures had been transitioned and their staff had been qualified and licensed to these new EOPs. They were in fact symptom-based, function-oriented EOPs.

In terms of the crew staffing compliments, at Loviisa, a normal sized crew represented four control room personnel, and the minimum crew represented three control room personnel for the study.

Next slide please. For the Halden study phase, we used a simulation of the Loviisa Nuclear Power Station process. So, the simulated plant at Halden was based upon the Loviisa Nuclear Power Station with added automation to simulate passive system performance.

Where we got the ideas for the automation were from the advanced reactor submittals. For example, Westinghouse had identified what the main differences were between the current generation Westinghouse and the future AP 600 in terms of passive system features. We tried to simulate those things in Halden through added automation, giving to the operators the look and feel of this passive system.

The other main feature about this simulated working environment was that the main control room at Halden was completely digital. The features that were selected for the main control room in Halden basically came from the advanced reactor Digital I&C submittals. So, it had a common process overview display, which is shown here in the middle, that both of the panel operators would share, which provided an overview of the process.

They each had a dedicated set of alarm displays that were digital. They had a set of process displays in selectable computers down here, selectable workstations, so they could bring up different parts of the plant. They could bring up different graphics for displaying information about the process and other selectable features.

Finally, in the center, they had a common safety parameter of display systems. This shows that portion of the laboratory that was configured for the reactor operator or the balance-of-plant operator. There was also, for the configurations in which there was a control room supervisor and a shift technical advisor, a set of displays back there for those people.

For the normal crew in the advanced plant setting, similar to at Loviisa, the normal crew configuration was four operators. The minimum crew at Halden was two operators. Because we have the need to maintain the same division of labor though, what that meant was that in some cases in the two-person configuration, one of the operators would be a dual role: operator/control room supervisor.

Next slide please. Observing that it's five past twelve, do you want me to continue?

MR. POWERS: You just go right ahead.

MR. HALLBERT: Okay.

MR. POWERS: I want to get this as a package.

MR. HALLBERT: All right.

Eight crews of licensed reactor operators and control room supervisors, senior reactor operators, participated in the study. Each crew experienced the five scenarios in different orders to handle counterbalancing effects. Four crews served in the normal and four crews served in the minimum staffing configurations.

MR. KRESS: Are these eight different crews?

MR. HALLBERT: These are eight different crews. And in the NUREG, it shows a layout of the experimental design as well too.

MR. POWERS: So I should see four data points on every plot, right?

MR. HALLBERT: Unless they're aggregated. That's true. Yes, that's what you'll see on some of the plots back here.

I'd like to talk a little about the social science underpinnings of the research now in terms of the data that we collected.

We collected data on a number of subjective performance measures. We were concerned, first and foremost, about changes in control room workload. In other words, the workload that the individual operators and the control room crew as a whole would experience as a result of changes in control room staffing. In other words, if the plants are fundamentally different and there are fewer things for control room operators to do, then you would expect to see differences in workload. So, we looked at workload and we measured that using the NASA Taskload Index measurement technique.

We were also interested fundamentally in what would happen to team performance. What we mean by that is what would happen to leadership characteristics in the main control room. What would happen to communication with fewer people and the same demands? What would happen to the focus on the task and the mitigation activities at hand, the esprit-de corps and things like that?

So, there was a measure technique called BARS, which is an acronym for the Behaviorally Anchored Rating Scales. That's also described in the NUREG as are all of these. That measurement technique taps into these team interactions.

Finally, we were also interested in the subjective measure of situation awareness. You've probably heard situation awareness discussed in the aviation industry quite a bit. That's where it was originally studied. What situation awareness refers to is primarily how well an operator understands what's going on around him or her in the plant.

MR. POWERS: Situation awareness is something the committee is fairly familiar with because it's a primary thing in the power upgrade issues.

MR. HALLBERT: Yes, it's very important.

There has been considerable research showing linkage between situation awareness and performance in the aviation industries. So, we had a measurement technique that was developed specifically to measure control room operator situation awareness. We were interested in what would happen there.

I want to also say that prior to this time, there really hadn't been any data collected on these kinds of measures in control room crews. So part of the study was also to gather a baseline of data of what happens to situation awareness, workload, and team performance during these kinds of scenarios. Not just under the study, but what happens to these things during the course of a transient.

We were also interested in objective performance, how well the crews managed the burdens of announcements, notifications, communications, for example, how well they perform their critical mitigation activities, and how well they've managed the longer activities of stabilization and cool down of the plant. These scenarios were obviously fairly long, ranging from an hour and a half to two hours in length. So, we looked at objective performance measures as well.

MR. POWERS: Let me ask you a question. You ran a scenario for an hour and a half?

MR. HALLBERT: Yes.

MR. POWERS: You say, "Okay team, we're going to start", run it, and then they know when it's over. But the reality is a control room operator sits there for an average of six hours and then there's an event and it's over.

MR. HALLBERT: Yes.

MR. POWERS: How does that factor that you don't have that lead in six hours affect performance?

MR. HALLBERT: Well, we knew for example that bringing these crews into the simulator with us foreign staff there was going to raise some expectancy on their part, so we told them what we were doing. We had a briefing package. That was necessary not only for this kind of research, but it was necessary for informed consent.

But what we did to address that concern was that all these scenarios typically began with a period of normal activity. We didn't want them to be conditioned to the fact that 15 minutes after we start this scenario, there's going to be something go wrong.

So these normal periods, for example, were load following or a perched control rod or something going on with the balance of plant, some sort of normal evolution. But typically it would last anywhere from 15 minutes to an hour or so to try to get them to relax a little bit and off edge.

MR. ROSEN: Then the scenario would start?

MR. HALLBERT: Then the scenario would start, then the transient would be introduced.

MR. POWERS: But for some reason, you thought 15 minutes to an hour was enough to simulate six hours or nine hours or twelve hours?

MR. HALLBERT: We relied upon the training staff at Loviisa to guide us in that kind of determination. We asked them, how much is enough to try to get them off the edge of their seats, and try to memorize their displays.

MR. POWERS: Somebody must have looked at this because it's the same problem you have in simulators every place.

MR. HALLBERT: It's like for re-qualifications I imagine. I mean you come to a training simulator expecting to learn some new thing, but also you expect to be challenged I suppose. So yes, that was an issue.

Let me also mentioned that this is what's referred to as repeated measures, experimental design in the sense that we collected data on these measures up here throughout the scenario. I'll show that starting on the next slide.

Next slide please. This is the part where I'll have to get up here and talk a bit. I'm going to talk a little bit about the results now in terms of their basic value to the staffing study and also try to illustrate some of the connection points with the issues of human reliability.

I want to explain, starting off with this graph up here, what it refers to and what it represents. This graph is a plot of workload that was measured throughout the scenario. Across the bottom her you see the scenario periods. This is an average plot across all scenarios. It shows average workload as it was experienced by the operating crews in this entire study. It's a generalized or normative kind of graph of what happens to workload during these transients.

As I mentioning, during the first scenario period, crews were conducting some kind of normal activity, normal evolution in the control room together with staff in the plant. We simulated plan personnel outside the control room to make these scenarios very realistic as well too.

Between scenario period one and scenario period two, the transient or transients were introduced. Then the scenario progressed for the duration of the particular scenario at hand here, this period out here at number five being a representation of time at the end of the scenario.

MR. ROSEN: What is the "Y" axis, a percent or what are the units of it?

MR. HALLBERT: This is measured workload. The NASA TLX Inventory measures workload on a scale from 0 to 100. And so, this shows that the crews on the average, their average workload during normal operation was rated as 25 out of 100.

MR. ROSEN: Where the 100 would be like running around like ants in a hive, going as fast as they can in every direction?

MR. HALLBERT: Something like that. Yes, I'm sure that would be the highest workload you could imagine.

What we see in here is that workload increased substantially from baseline operations during the disturbed phase of this scenario. A couple of things are worth discussing about this graph here.

The first is that it shows what happens to operator workload during these transients. The second thing is that the National Research Council has studied for the Department of Defense the issue of workload transition and workload in general and have identified a couple of concerns.

MR. KRESS: How did you measure this

workload? I'm not familiar with this index. Did you have the measure of his metabolism rate or did you ask him how busy he was? Is this subjective?

MR. HALLBERT: It's subjective. There is a standardized technique, a standardized psychological measurement technique, and it's called the NASA TLX because NASA developed it. It refers to taskload index. And, there's a standard set of instructions and a standard form for measuring. The taskload index is also described in that NUREG. It's shown in the appendix in the back there.

MR. KRESS: Okay.

MR. HALLBERT: But it taps into a number of relevant workload factors such as temporal demand, physical demand, mental demand, and things like that.

MR. KRESS: But you or someone like you sat there and filled in the numbers?

MR. HALLBERT: No, we didn't. What we would do is -- and that's a good question because it gets to something that I glossed over in here. What we did was at certain phases of the scenario, we would pause the simulator and we would administer these instruments. Then the operators themselves would rate their workload during that scenario period. Good question.

MR. KRESS: For a given scenario, you probably could've decided on what the workload was ahead of time.

MR. HALLBERT: Well, there was a shadow study in this research, as Jay mentioned, on simulating operator performance. We developed a simulation of operator performance that predicted workloads using some predictive techniques. In general, those correlated well with the workload the operator experienced I think, if I recall that study.

Getting back to this performance curve here, the National Research Council identified two primary concerns of workload. One is it's acute effect in what they call workload transition. That's illustrated here in the change of workload from time period one to time period two. The concern is that during periods of workload transition, errors are likely.

The other concern that was identified and has been identified in the open psychological literature are the chronic effects of the workload. In other words, we know that experts such as licensed reactor operators are able to mask performance of a situation even under situations of high demand and high stress for a period of time. But that overtime, those high demands place burden and stress upon the humans in the control room and at some point out here, performance degradations are more likely.

MR. KRESS: Is this a linear time scale?

MR. HALLBERT: This is a linear time scale, yes.

MR. SIU: Each of those are equal time?

MR. HALLBERT: Pretty equal, yes.

MR. SIEBER: Did you measure error rate?

MR. HALLBERT: We did not in the study because -- the main purpose of this study was not to focus on the errors. It was focusing on performance and control rooms and trying to evaluate the issue of staffing. We did not study error per say.

MR. SIEBER: I would've thought that would've been a key element to decide what size crew you would apply to what kind of a reactor.

MR. HALLBERT: No, we --

MR. SIEBER: Because if you don't have enough operators, you're going to make a lot of mistakes.

MR. HALLBERT: No, we didn't.

But, we measured something else, which was their performance in mitigating the transients. What we believed was that their ability to manage all the responsibilities in the control room including announcements, notifications, activations of fire departments, emergency operations centers, all those kinds of things, would eventually show up as an effective reduce in the crew size. The hypothesis being that crews with a normal size would be able to be better managed objective performance than smaller crew size, all other things being equal. But, we know they weren't because there was also automation and passivity in the advanced plants.

MR. SIEBER: Thank you.

MR. HALLBERT: I'd like to talk now about the other subjective performance here on the graph, which was situation awareness.

Up to this point, we hadn't really had a good baseline of measurement of situation awareness on control room operators. What we found was that similar to the graph here for workload, compliments sort of occurred or the reverse sort of occurred to situation awareness. As workload was going up, situation awareness was going down.

MR. POWERS: What I don't quite understand on all these plots is if four crews do this --

MR. HALLBERT: Eight crews all together.

MR. POWERS: Right.

MR. HALLBERT: Representing something like 40 operators or something like that.

MR. POWERS: Each one of those points should have unless -- I mean the remarkable thing, everybody was identical here. I can't imagine.

MR. HALLBERT: This is averaged.

MR. POWERS: If it's averaged, then can you give me some idea of what the variance was in that average?

MR. HALLBERT: Yes, there were a number of interesting findings about the variance itself, which is almost the subject of a separate discussion.

In fact, that is shown in the report. There were significant variations in situation awareness as a function of conventional verses advanced and minimal verses normal crew staff and sizes. There were some significant variations there that contributed to the main findings.

MR. POWERS: If I go to interrupt these results, what do I communicate to the HRA folks about this? Do I just give them the means or do I use the means and the variance to compute 95 percentiles or something like that? I mean what number do I actually use?

MR. HALLBERT: I think, if you're asking

me if I were communicating to another HRA person, -- and I consider myself to be an HRA person -- I would say when I look at these results, I see some general trends that are relevant during a scenario. And that is that, after the onset of a scenario, the crews are required to make decisions that require high degrees of situation awareness. If there was a higher degree of likelihood in making those decisions or making a decision, they're at greater risk for an error.

The other thing is that even though the recovery of situation awareness approximates its loss, the recovery is invariant. Factors at the end of the scenario are factors that the crews in fact themselves introduce. So, we weren't doing things out here. The manipulations we made to the scenarios typically ended somewhere right around in here or so.

MR. POWERS: Right.

MR. HALLBERT: So, losses in situation awareness here were not due to anything that we had done. These were due to things that the crews had done themselves. So they, in some way, lost control of the situation maybe to some respect and didn't have good situation awareness at the end of the scenario. And, there are still critical decisions out there to be made.

The other thing that I would say, and it gets into the subject of PSS, there are some important scenario specific differences in situation awareness. I don't have a graphing here. It's in the NUREG. But, we did find differences in situation awareness between what we described as rule-based scenarios verses knowledge-based scenarios. That's using a term coined by Jens Rasmussen, a researcher in this area.

What he posited, that process control was achieved through a variety of different situations based upon the degree to which they were readily established rules available for operators to follow such as procedures, matching the situation exactly verses situations in which a high degree of interpretation was required on how to apply those procedures, being more of a knowledge-based kind of a scenario and other things like that.

MR. ROSEN: Now you've got me confused because you told us earlier that these operators were using symptom-oriented procedures.

MR. HALLBERT: That's correct.

MR. ROSEN: Which you do not need to know the situation in great detail at least early on.

MR. HALLBERT: You don't require diagnosis to select the appropriate final procedure. In other words, you can maintain your critical safety functions using the symptom-based procedures. But eventually, for every procedure, you have to transition out to the appropriate -- what's it called, recovery procedure?

MR. ROSEN: No contest. I agree with you.

MR. HALLBERT: Okay.

MR. ROSEN: But in the early phases, maybe on the left hand side of your curve, situation awareness is not all that important. He's following his symptom-oriented procedures. He looks at the symptoms and takes the actions that the symptoms require.

MR. HALLBERT: There may be some decisions required early in a scenario as to what systems to use and in what ways depending upon the ways in which systems fail.

I'll use an example of a loss of feed water. You lost the main feed water pumps and now you have to use your auxiliary feed water system. Well, if there are certain malfunctions or certain systems out of service that complicate that decision, you do have to have good situation awareness in order to make a decision about how to recover those systems.

MR. KRESS: This point 7, is it good or bad awareness? Is it an A, B, C, or D?

MR. HALLBERT: I don't know. I'll be honest with you, this was the first time that we have collected data on these kinds of performance metrics. We did it for the purpose of this specific study, but I don't know that we really know how much situation awareness is enough.

What I can tell you though, is that when you get down to levels of point 5 that's situation awareness. And that means that your ability to understand what's going on in the plant with regard to all your systems is about half right and about half wrong.

MR. KRESS: Fifty-fifty chance.

MR. HALLBERT: Fifty-fifty. And when you start dropping below that, there are some --

MR. APOSTOLAKIS: Overall, did all the crews exhibit specific behavior?

MR. HALLBERT: Overall on an average, the answer is "yes". This is the average. The specific question, did every crew experience it this way? I would have to go back and look at that data, George. There were some transient specific differences like I said.

MR. WALLIS: All this is fascinating but I don't know what it has to do with regulating nuclear reactors. It's very interesting but I don't know what to make of it if there's no hypothesis being tested or anything.

MR. HALLBERT: The particular issue under study here was what would happen to control room crew performance if you were to make changes to main control room staffing as well as made a transition to these advanced reactors.

Our purpose in conducting this was to provide a technical basis to the Office of Research to supply to NRR in making decisions about what information would you require of a licensee to show that performance was adequate in this new situation, as an example.

MR. WALLIS: This must be dependent on all kinds of things, all kinds of scenarios, or all kinds of stuff. So to get anything as generalized as a lot of this must be very difficult unless you have a big database or some good hypotheses or something.

MR. HALLBERT: Well, in terms of the actual reference values for how much situation awareness you need to have in a new system, you're right. We don't have that number yet. We haven't published it. We haven't really even thought about it. But in terms of looking at the implications of this research, there was some generalized findings from it. Again, that's described in the NUREG.

My point here was to try to show that in this research there were some connection points between operator performance and the general issue of human reliability. That being that there are situations in here in which performance will degrade. And those situations can be studied to extract information.

Next slide. Another question we had was how well do these performance methods, the subjective performance metrics correlate with their objective performance. So we looked in a few areas and here, Dana, is one of your eight point graphs that you were saying you would expect.

What we found in one set of analyses when we measured team performance, how well they communicated/interacted as a crew, the trends there paralleled their objective performance in managing the transient. So indeed, that factor of team performance appears to be a vital one for controlling and managing the transients. We found that also out in the study here.

Again, the implication being for HRA, that if you start doing things that affect the ways the crews work, like you were talking about earlier, people and crews that don't normally perform in crews and things like that, those implications need to be thought of because there may be attendant affects on preponderability to manage these kinds of transients.

Next slide.

MR. SIEBER: One second.

MR. HALLBERT: Yes.

MR. SIEBER: Both of those plots cross, and it appears that in the advanced plant you're better off with a smaller crew.

MR. HALLBERT: There were some significant interactions in the study here. What we found was that, in this particular case, the minimum crew in the conventional plant did not perform as well as a normal size crew.

If you could imagine, for example, in a normal size plant, it's designed for a larger sized crew. When you go to an advanced plant that has a more compact control room and it has more design features for a small sized crew, their performance was as good as the normal size crew and better in many instances.

Next slide please. I'm going to talk now about the embedded study that was carried out within this larger study on control room staffing in advanced plants. I talked about it earlier.

The intent here was to collect data on operator performance and performance shaping factors. Performance shaping factors, as most of you are probably familiar, is a term and concept that's used frequently in many human reliability analysis methods. The way that it's often used is that there's often times a nominal or assumed human error probability for a certain kind of action, and that nominal human error probability is modified for the effects of certain performance shaping factors. This includes things such as training procedures, human machine interface experience, and things like that of the crew.

So, there is and always has been for as long as these two concepts have been around, some intuitive linkage between performance shaping factors and operator performance. I think Alan Swain described the linkage very well in NUREG 1472.

As a whole, the types of PSFs and their affects on error rates vary quite significantly among the HRA methods that are out there. If you look at them, you'll see that the effect on HEPs vary significantly.

The way that these effects are assessed is currently that they are estimated. Analysts or a group of analysts will sit down and will say: how much credit do we give the operators for having good procedures in this scenario, or how much do we credit them for their experience and training. As a result, there is a fair amount of uncertainty really in the effects of these PSFs on human error probability.

So, my belief was that there was a need for a better benchmarking and understanding of performance shaping factors with actual performance. And if we had that linkage, we could build better models of failure eventually.

Next slide. So the purpose of collecting data about these performance shaping factors was to explore how these things could support HRA, these larger human factors studies.

The specific objectives were to identify a set of performance shaping factors that were predictive of crew performance, determine the relative weighting of these factors to one another, develop or demonstrate a general model in which these performance shaping factors could be expressed one to another with operator performance, measure the factors affecting the predictive validity of these performance shaping factors, and replicate the results.

MR. KRESS: The performance shaping factors were independent variables in this study. Were they varied one at a time or several at a time?

MR. HALLBERT: No, I didn't do that. In fact, what I was essentially doing was piggybacking on the previous study that I mentioned. So, I took the performance shaping factors --

MR. KRESS: I see. You took exactly what was in there?

MR. HALLBERT: Yes, exactly how they came. There were some good things to that and there were some bad things to that. We can discuss that.

MR. KRESS: It relates to how you design experiments?

MR. HALLBERT: Exactly. I mean ideally, you'd like to measure one at a time then add a second and maybe a third then maybe a fourth. But the counterargument to that is you never have just one or two or three. You have them all. So, I took them all because that's what I had and that's what I was given.

Next slide please. This research really started back in the middle 1980s when we had the opportunity to collect data on performance shaping factors as part of other studies. I mentioned NUREG Contractor Report 4966. That's where that work was originally documented.

At the time, we developed an instrument to measure performance shaping factors' affect upon operator performance. Through analyses and reductions in data, we identified that really seven of these ten performance shaping factors really had some predictive power, and that the other three really didn't seem to matter to the crews.

MR. KRESS: What were the other three?

MR. HALLBERT: There in 4966, but I don't recall them. Maybe even the way that they were defined was vague. Not that they didn't have an effect, but the way that we had defined them could've been unclear to the crews.

The ones that did have effects and were demonstrated through statistical analysis techniques included aspects of procedures, training, stress, workload, information available to the crew, the way that the system provided feedback to the crew on their actions, and the human machine interface in general.

MR. KRESS: Is time required to do an action? Is that a performance shaping factor or is that something else?

MR. HALLBERT: That was actually the dependent measure.

MR. KRESS: It's a dependent measure?

MR. HALLBERT: Yes.

MR. KRESS: How long it took him to actually do the --

MR. HALLBERT: The important thing that they had to do in that particular scenario was what we actually measured. I'll explain how we did this just a bit more here now.

We had a data collection instrument that we developed to measure how the operators experienced these performance shaping factors. In their own terms, how they affected their ability to carry out the critical mitigation tasks in a particular scenario. We asked them to rate these performance shaping factors just after the completion of a transient, a scenario study if you will.

MR. KRESS: The instrument could be a form that they fill out?

MR. HALLBERT: It was a form. That's exactly what it was.

We asked them to consider each of these performance shaping factors that we had discussed and defined prior to their running the scenario. Then we afterwards asked them to rate on a scale how these things had influenced their ability to take the appropriate mitigation action, which was specifically defined.

In the case, for example, the loss of feed water, it was to restore the condensate booster pumps. In the case of the LOCA, it was to isolate the hot lag or something like that in a particular scenario.

After the simulator trials were done, these operators rated the affects of the PSFs on their performance of the critical mitigation tasks. The data that I'm going to present today is essentially the result of collecting data at different times with different crews and different locations.

We had four crews in the US plant and that's documented in this NUREG reference here. We had four crews at Loviisa like I was just describing, and then four crews at Halden. And, we had three common scenarios: undercooling, overcooling, and a loss of coolant scenario. Again, we had the thermalhydraulic references for all these scenarios. We thought they were comparable in nature.

Next slide please. The results are that we used a linear model to assess the effects of the performance shaping factors on operator performance. Whereas I mentioned previously, the prediction of "Y" in this formula here was the critical task mitigation performance time. When, after the initiation of the scenario, were they able to complete their critical mitigation task?

We collected data on these performance shaping factors across these scenarios, crews, and plants, and even countries I suppose. What we found --

MR. POWERS: What does it mean when you use a linear model like that with a constant term? It becomes an adjustable parameter in this model.

MR. HALLBERT: It actually was empirically driving. What we found was that -- and you'll be able to see on the next graph, the next slide -- that typically the prediction of performance would intersect with the "Y" axis, and the effects of these performance shaping factors were over and above, or were around, that intersection point.

So let's say, for example, that the average mitigation time was 18 minutes after the initiation of the scenario. You could have the intersection point being at 14 or 12 minutes. Then the PSFs basically predicted up and around -- or the weighting of these factors predicted up and around that time.

What we found through these studies and

the data collection was that the linear model was sensitive to scenario differences. And I'll show you on another slide how we found that. It was sensitive to plant differences, and it also demonstrated predictive ability.

Next slide. I talked about being sensitive to plant differences. Here is the sum total aggregation of the normalized critical mitigation times. These are the predicted values.

We see, overall, that the multiple correlation in the multiple regression model here was 0.36. What that means is that about 14 percent of the variability in the scatter of the actual mitigation time can be predicted by that model.

MR. WALLIS: Now it's predicted based on data? It isn't a prediction from something else?

MR. HALLBERT: It's a prediction from the best fit of that linear model.

MR. WALLIS: So when you have a limited amount of data and a number of coefficients, you're going to predict something even if it's --

MR. POWERS: What he actually is looking at is what fraction of the variance in the data can be explained with this linear model?

MR. HALLBERT: And the unique contribution of the individual performance shaping factors is measured through the beta weights.

So, this is all crews, all plants, all scenarios. Fourteen percent of the variability was explained through this linear model.

When you looked at just one plant for example, all scenarios, the multiple correlation coefficients were significantly higher. And, you found the same result for all the other plants. So what you see is that the predicted model has greater predictive ability when looking at specific scenarios as opposed to all scenarios. We went from explaining 14 percent of the variability up to about 47 percent of the variability.

We found the same thing in plants. In other words, the closer you got to specific scenarios within a plant, the greater the predictive ability of the model was. So this is suggesting something. It's suggesting that individual differences and how operators experience the scenarios is significant. They are truly different. For example, an implication of this might be that how would we recommend people incorporate performance shaping factors into a particular scenario.

Next slide.

MR. POWERS: What you're really saying is that there's not a uniform PSF for every scenario, that I just can't put a constant in there?

MR. HALLBERT: That's right, and it seems to be different across plants.

MR. ROSEN: It doesn't seem to be that surprising, does it? That operators would react differently to undercooling than they would to overcooling, that they would react differently to loss of power? But within those three scenarios, that operators would feel more challenged by undercooling for instance.

MR. HALLBERT: Or more along the lines of what aspects of their procedures and training and other performance shaping factors contributed to their ability to mitigate that transient, and how then in the future to best incorporate those performance shaping factors into the estimation of human error probabilities.

Again, this is part of establishing a technical basis for how performance shaping factors should be addressed in an HRA.

MR. POWERS: Yes. I mean that's what he is really -- he hasn't got anything definitive here, but he's building an information base that's really calling into question the way we do things now. As George said, we do get smarter with time. It's not always obvious we get smarter. All it says is life is more complicated than we thought.

MR. APOSTOLAKIS: Let's not be unfair. People do consider different performance shaping factors for different scenarios in existing models. And, it's nice to have confirmation of --

MR. POWERS: But see, what he's saying is that if you take a specific performance shaping factor and say it's affect is to double the time, that may be true for one scenario, but it may not be true for another scenario.

MR. SIU: That's right. Some HRA methods, indeed, they do allow you to adjust and others they don't. Now for guidance, it raises immediate questions.

MR. APOSTOLAKIS: I mean you see more clearly that --

MR. POWER: More pertinent is that he's also demonstrating that you can actually get something useful out of these studies, which is really excited.

MR. APOSTOLAKIS: I don't think anybody else has done this, have they?

MR. HALLBERT: No, not anything like this.

I kind of combed the literature. Again, the reason why it's been sort of a passion of mine over the years has been because there is such apposity of information about these things. The other things is that it really is needed I believe.

MR. APOSTOLAKIS: So all we needed was a passionate guy.

(Laughter.)

MR. POWERS: That's what's needed in everything. I mean if you hadn't had runners cruising down the mile, we would not understand anything about the momentum of the equation.

MR. HALLBERT: I'm actually more passionate about other things, but this is very interesting.

The other thing that I wanted to mention is that there would be some intrinsic value to not only understanding about the performance shaping factors' relationship on performance, but for example, how important certain of these performance shaping factors are in certain kinds of scenarios. Now I haven't done that analysis yet. I'm interested in looking at it, but I haven't done it yet.

For example, we talked about: are there any properties that are unique to undercooling scenarios that are demonstrated through these performance shaping factors. I don't know. I don't know yet.

MR. APOSTOLAKIS: So now you're really creating the context within which the HRA modeler would develop the models, the general shape of the models. I think this is great.

MR. HALLBERT: Yes, hopefully. And even eventually to provide some insights and better guidance.

MR. POWERS: To be precise George, the context with which they will evaluate the plethora of models, we'll see if they're useful or not.

(Laughter.)

MR. HALLBERT: And perhaps even from a regulatory perspective, eventually to be able to asses the HRAs that are done and to find out whether all the appropriate PSFs have been taken into account.

MR. POWERS: Yes.

MR. HALLBERT: And why they believe so or not.

MR. POWERS: But let us not forget, if you're seeing -- this is not unusual in this stage of understanding to have a substantial amount of the variance that remain unexplained.

MR. HALLBERT: Yes.

MR. POWERS: It's terrible, but I happen to know in a lot of physical fields that that's where we start, with huge amounts of variance and discovering where that increment of variance is --

MR. HALLBERT: Yes. I mean this to me is very exciting because what you're describing is very applicable to this stage right here. There has not been a lot of data collection yet and it's very informative.

MR. POWERS: From a statistical point of view, the problem with your model and your procedures is that what you're treating as well known variables for themselves have a substantial amount of uncertainty in there, and you've used a liner regression analysis in which you're assuming that those things are all precise. You shouldn't have done that. But unfortunately, the regression algorithms for the right way to do that are pretty hairy to work with.

MR. HALLBERT: Yes, and also in the social sciences, these liner regression models have been shown to be fairly robust to certain violations of assumptions and mathematical properties. So, we start there and at least try to establish that there is a relationship and try to understand better the appropriate models eventually.

MR. POWERS: Try a min-max routine against this and see if it doesn't give you -- first of all, it'll eliminate a certain amount of your variance.

MR. HALLBERT: Yes, min-max or stepwise approaches. Good recommendation.

MR. APOSTOLAKIS: You should keep everything in context. You're not producing a --

MR. POWERS: He's looking for a variance that can be explained and what not. Now some of his variance comes from the fact that his independent variables are just themselves uncertain.

MR. HALLBERT: Thank you. I'll summarize now the presentations of both the embedded study and the overall point of my presentation.

First of all, in the embedded study --

MR. APOSTOLAKIS: Excuse me. Can you tell also at some point what is the most important performance shaping factor or the top three?

MR. HALLBERT: I hate to answer your question this way, but we did some exploratory analysis into the relationships among the performance shaping factors, and we found some stability through factor analytic reduction techniques. Essentially, you could sort of define three overarching performance shaping factors in the set of seven if you will. Stress and workload basically comprised one factor that we'll call demand or maybe even workload. But they loaded negatively overall in these scenarios. So what they did was they kind of worked against the operator.

The other ones were procedures and training, procedures and training loaded together. And, that seemed to be best described as preparedness, how well prepared they were to deal with the specific demands of the transient.

The other three were information available, system feedback, and the HMI, which is probably best described as the HMI. So, features of the control room design, features of the crews' preparedness, and the control room systems designed for the scenarios, as well as the crews own experience of the transient and it's negative effect upon their ability to match with the demands.

MR. APOSTOLAKIS: But is available time and performance --

MR. HALLBERT: I didn't define -

MR. POWERS: He has taken that out of his study because that's what he's measuring in "Y".

MR. HALLBERT: I agree with you.

MR. APOSTOLAKIS: If I have a task that needs to be completed in 20 minutes verses another one that's 42 minutes, should I asses the impact of the time difference on these preparedness performance shaping factors and then do my analysis, or do I have guidance as to how the 20 minutes verses the 42 minutes will effect it? Should I go indirectly through the three that you mentioned or it is from the factor itself?

MR. HALLBERT: I don't know.

MR. APOSTOLAKIS: Again, I don't expect you to have all the answers. But, these are the kinds of questions I think that are important.

MR. HALLBERT: It's a limitation of the approach that --

MR. POWERS: The way that he has done his study, he can't really answer the question.

MR. APOSTOLAKIS: That's fine.

MR. POWERS: He didn't say you were wrong. It just said, I have to look at a --

MR. APOSTOLAKIS: He would never say that even if he thought it.

MR. POWERS: We will say that for him.

(Laughter.)

MR. POWERS: I think we can move on.

MR. APOSTOLAKIS: These things are things that we ultimately have to face in certain regulatory actions.

MR. WALLIS: We have faced already. We have some data.

MR. POWERS: I think we can congratulate you on a pretty well defined study. I can quibble with your data reduction techniques, but I know what you're trying to do. I think it's interesting that you're getting insights out of this thing, which is all you can ask for right now. The actual percentages, that will have to come with time.

MR. HALLBERT: Yes.

MR. POWERS: I think we can -- unless you have some particular points you want to make here.

MR. HALLBERT: The think the final slide was just essentially what I've already covered. To date, there have been some studies conducted and there is some data available right now. And, we're looking through those sources of data to see what is relevant for HRA.

New studies offer great promise because whatever we learn from these other studies could be taken into account for the design of future studies to collect specifically, HRA -- or these kinds of questions from the outset that you're asking us today.

Then I kind of end up where I started, which is that I believe these simulator studies are valuable, and they provide useful data for HRA.

MR. POWERS: I would put a caveat on that. I think simulator studies carefully designed, well conceived, appropriately done, and cautiously used can yield insights that perhaps give us an idea on what we ought to be doing.

MR. ROSEN: Just like thermalhydraulics studies.

MR. HALLBERT: I agree with those points you just made.

MR. POWERS: I mean I think that's the step that this committee has never seen, people coming in and doing simulator studies very carefully, very well designed with particular objectives. They may well have done that, but we just have never seen it.

MR. APOSTOLAKIS: They keep it a secret.

MR. POWERS: Well, there's always a problem when you present to this committee that doesn't pretend to be specialists in this field. But this was nice. You could understand it and what not.

What I would like to do now is quickly ask the members what kinds of topics they want to pursue further in the discussions this afternoon. I think we're done -- am I correct in thinking that we're largely done with the formal presentations and now we want to discuss what the research program is going to be?

I myself very much want to go into this topic that showed up on both Erasmia's slide and Jay's slide called tools and tool development. I'd like to understand what the objectives of tools are, what the vision is, who those tools are for, what they're going to look like. And I invite the other members to make comments on what they want to talk about when we come back from lunch.

MR. ROSEN: I'd like to talk about the issues of organizational performance, safety culture, and indicators.

MR. APOSTOLAKIS: Seconded. Also, in addition to this, I would like to understand a little better the development of the plants to develop an HRA model that will actually give distribution. I mean is there a conceptual design at this point or that kind of thing? I know that it's still early.

MR. POWERS: You get the chance to name your topic, not discuss your topic right now.

Graham, are there any other issues that you'd like to pursue?

MR. WALLIS: Nothing more than yours that really asked the questions so far. I'm an interloper on this committee anyway.

MR. POWERS: You are never an interloper. You are a very welcomed participant.

MR. WALLIS: This is a very tough area to quantify. It's much tougher from the hydraulic. And I don't quite know what tools could be useful and how they would be validated. So, I've asked questions like that already.

MR. POWERS: Dr. Kress?

MR. KRESS: No. I'm interested in it too.

MR. POWERS: Okay. Jay?

MR. PERSENSKY: I'm also interested in tools and safety culture issues.

MR. POWERS: George?

MR. APOSTOLAKIS: Can I add one more?

MR. POWERS: Yes, you are unlimited to the topics.

MR. APOSTOLAKIS: The view of existing models and what plans there are to use them in the development of your own model would be of great interest to me.

MR. WALLIS: What's the state of that?

MR. APOSTOLAKIS: Yes.

MR. WALLIS: You had those four operators in so many situations or something. Now, that's an interesting study. But there must have been a lot of things like those before in some other context.

MR. POWERS: Well, as far as care of design, this is one of the best I've ever seen.

MR. WALLIS: Like human performance in flying airplanes.

MR. POWERS: Now let me interrogate our speakers. What would you guys like to talk about this afternoon?

MR. SIU: Actually before we get to that, I think one point to make is that Bruce and Dave Gertman have a flight and they to leave here by about three o'clock. So, any questions that you have relating to I think the last point -- well, I guess you'll obviously have something to say about existing models, but also if you wanted to talk about experiments that would be good right after lunch to make sure those get done.

MR. POWERS: Okay, the experiments right after lunch.

Are there topics that you need to

communicate? Recognize that our intention is to write a letter that says here are the aspects of the research program that we like and what not. So if there are things that you think we need to understand better, don't be shy about it.

MR. WALLIS: I have a question. We had some very general presentations about the program then we had something very specific from Bruce. There must be other specifics that are going on that would illustrate the generalities for me.

MS. LOIS: So then the intent was to give you an overview of where the program has --

MR. WALLIS: But it seemed to be that we went from one pole to the other.

MS. LOIS: But we hope that this will be the beginning of probably several follow up meetings with the committee to tell them in more detail. On the things that we've done in detail -- I guess those that are still in the planning stage, we're just struggling with that, some things.

MR. WALLIS: In the case of thermalhydraulics, we have some sort of big scheme of needs and then we have the framework, which is codes, and then we have individual projects fitting into the codes. And because of an individual project, we've got some kind of general scheme. What the code of mechanism that --

MR. POWERS: The grand vision is what you're talking about, and that's where I want to go with the tool development and try to understand that a little better in the grand scheme.

I think we are going to get an opportunity to see the applications that showed up frequently. I'm much more concerned right now about the underlying technology we're developing that supports all these applications, the PTS, and things like that that are going on, and the strength of that program. And, we'll discuss that.

In that case, I propose we go ahead and break until 2:00.

(Whereupon, the committee recessed for lunch at 1:00 p.m.)

A-F-T-E-R-N-O-O-N S-E-S-S-I-O-N

(2:02 p.m.)

MR. POWERS: Let's come back into session.

We concluded the last section by saying here are the things that we want to talk about. It looks to me like the topics, the big scheme of the program, what we mean by tools, organization, safety culture, indicators, development of HRA models, and the view of existing models and the state of the art are the topics.

It does not look like we are going to go into any great detail further on data collection and data manipulation and digest. Though, I will emphasize to you the concluding talk on which we did there was illuminating and gives us new insights on the importance of various elements in the program book, the human factors and HRA.

At this point, I'd like to understand better the program, what's in place, what's in just the planning stages, what we're trying to endorse here exactly. Okay?

MR. SIU: Let me start off by saying that we've asked John Forester and Dave Gertman to join us at the side table. I hope they'll chime in with comments as the discussion moves along. Of course, Bruce Hallbert is sitting up front with us. And Dave and Bruce, again, have to leave at about three o'clock.

A number of questions have come up about the vision of the program. I guess I'd like to get to that a little bit. We tried this morning to give you some sense of how we saw things. Obviously, it wasn't detailed at all and it wasn't intended to be.

Let me start by saying that I think that there are two aspects of vision. One is, if you will, organizational, and one is technical. The organizational vision is pretty much what you were seeing this morning. We have needs presented to us from other parts of the agency. From our understanding of what's going on in other parts of the agency, we try to our best to help address those needs through the activities that we perform, which are analyses, reviews, and developmental activities.

This seems trivial, but actually it's not because this is one of the areas where we got good comments from NRR in their review of our research plan. They talked about the need for much more focus or emphasis on issues like HRA guidance. We had it in our original plan, but we hadn't perhaps put sufficient emphasis on that. So, this is one place where we think we're going to strengthen, to support folks who are faced with particular applications.

MR. POWERS: What's your view? It seems to me that there are two models for the support that you could provide to the non-specialist in this area that has needs.

One is that you can say, "Here is my telephone. Anytime you need an HRA analysis, give me a call and we'll get it done for you." Clearly, that's the mode you operating in now and it may well be the mode you have to operate in.

The other vision is to say I'll live that way for a while, but eventually I want to have tools in these guys hands so when they have an HRA question, they can pull up this tool that will act like an expert system, it'll walk them through the questions, and they'll get their own answers.

MR. SIU: I don't know that we actually fit into either model right now. I think what we would like to do is more towards the second. Where we are right now is actually, in the case of reactor operations, NRR is doing the HRA reviews. We are not doing HRA reviews.

What we haven't done, and NRR pointed this out, is we haven't taken the results of our research over the years and boiled it down to something that -- for example, a review of the use when looking at an application. And by "use", it doesn't mean necessarily redoing the analysis. You might just say, "What are the questions I should be asking?" These are things that I think in the short-term we need to be working towards.

MR. CRONENBERG: This morning the power uprates came up as an issue that the PRAs are coming in saying that there's no effect on human performance or little effect on the power uprates. Yet, they have the study where one of the principle impacts was the reduced operator time for reaction to accident scenarios.

And so, we had the conflict there on one -- it was a study, and then the licensees come in and say there is no effect, and this committee had to struggle with these types of issues in the last year and a half on power upgrades.

Have you had any user needs from NRR to answer questions like that or have you given them any support? They are not risk informed, licensed amendment requests. They are traditional licensed amendment requests, so risk information is kind of supplemental to those requests.

But still, this committee has struggled with conflicting -- and their gut feeling is conflicting with what the licensee is telling them.

MR. SIU: And quite literally, we do not have the user need to provide support there, at least in the HRA realm.

Jay, I don't if you guys have been?

MR. PERSENSKY: Not specifically to that. I mean the work that we were doing on the changes to the operator action was in fact in part related to the power uprates. In that, if it is a risk informed submittal, there is a way of dealing with the risk aspect of it. If it's not, we can still apply risk to it. But the basis there was more to look at the level of review.

As I understand it -- Dick was here earlier, and he's been one of the people that I know involved in that from NRR. Most of what they've been looking at for the power uprates, they've actually looked at simulator trials and requal trials and they found that the actual error rates, not HEPs, but error rates have been very low in that kind of a situation. So, they've been basing their approvals on that.

I just saw Dick walk in if you want to

follow up anything on that.

MR. POWERS: Well, let me follow it up with a question here. In the course of this morning's discussion, we had a variety of questions raised on the adequacy of simulator data to reflect what goes on in the actual plant. How does one take those questions and look at the simulator trials with a jaundiced eye?

MR. PERSENSKY: Some of the things that you indicated were problems. For instance, bringing in different people. Just like any other experimental situation, especially when you're dealing with people, you can do a very large, multi-variant experiment, but the time and resources and ability to do that is very limited.

From the standpoint of the situation that we're talking about here for the uprates, it's their plant, it's their operators operating primarily in their mode of operation rather than separate modes of operation. It's their normal mode. So that's what we asked them to demonstrate. The whole point is being able to demonstrate that they can do it with sufficient cushion I believe.

MR. POWERS: The question I'm not asking is, it's not a question of really power uprates. The question is one of research and what Nathan said. What kind of questions should we be arming these guys to ask when they look at that information, and what are we doing to develop those kind of questions?

Like I said, we came up with some questions on the fidelity of simulators for actual plant operations. I mean they're kind of a qualitative sense so it be difficult to defend that as proof. You just couldn't use that information at all. It was just totally inapplicable based on the discussions we've had, but it's enough of a question that shouldn't the research program be addressing that kind of question?

MR. SIU: Yes. And again, I think that was the intent of the guidance task in various areas. We would start relatively modestly in terms of taking what we've learned to date and then trying to if not make a formal guidance, at least provide some useful information to users. And later on, of course, start getting more formal in terms of guidance for specific things.

Erasmia had mentioned the HRA standards activity, for example, and we intend to play a more active role in that activity.

MR. ROSEN: To refer to that comment that

I made earlier was to me, the way you would handle that properly was it's just another performance shaping factor. It's a crew performance shaping factor. Whatever number you ascribe to the likelihood that the crew performs successfully as it is constituted normally, you modify that number with some shaping factor. But a third of the time, the crew is not going to be in its normal configuration.

MR. SIU: And research again, whichever way they answer laws could provide a basis for deciding when you can take a certain degree of credit or under what conditions you can take a certain degree of credit.

MR. POWERS: I think, I mean we've had licensees, or in this case the applicants, come in and say we go through THERP on this thing and we get 1/100, but when we look at our database we see it's more like 1/1000. Could we go ahead and use 1/100 to cover this? And Professor Wallis says, "How do you know that factor 10 is good enough?" That's the question that's really answered here in this guidance program. The other guys he has downgraded his information by a factor of 10. Yes, that's probably more than enough or it's half of what he should've been or something like that.

I mean that's what you mean by "guidance", how to ask a question and what kind of answer is a reasonable answer. It will never be out to two significant digits because every plant is different and every environment is different and what they can tolerate is different.

MR. FLACA: If I could just follow up with a comment on that. When we look at a number though, it really represents something. What's behind the number, of course, is what's important: the procedures, the framing, and so on, how likely the event is going to occur, and what the operator is going to be prepared for. So, I think it really represents the way one thinks about it. I think that's what George was saying before.

And the question is, as far as our programs are concerned, do we have the infrastructure to be able to think about these questions, and be able to answer other questions that might evolve from the pursuit of these changes that are going on out there? Whether we have the tools and ability to do that I think is very critical. If we don't have them, we're only kidding ourselves. We're just not asking the right questions. We don't know if we've got the right answers.

But in all that context, I think it's more than just a quantification. It's really looking at what that means in the context of what you're giving credit for. If it's 1 in a 100, we expected than there should be a certain level of backup, a technical basis for that 1 in a 100. That comes down to doing some analysis based on what procedures and so on is in place. And, we need the tools to do that.

Now the question I guess is do we need certain tools, do we need to develop new to come and address new issues? One of them is the changes in risk as we see them as plants are making changes. Some of this is maybe due to manual actions verses automatic actions or changing things in that way. And how do we go about doing that, and do we have the tools in place to do that?

Isn't that really the issue on the tools? Again, I'm sorry. I came in a little late and I didn't really hear the beginning of it.

MR. POWERS: Well, the issue in tools is -- you certainly hit upon an important aspect on the issue of tools. My particular interest is one of vision of what the tools we want to look for -- not in the next three years, but say in ten years - when we actually get advanced plants coming in here to be certified, what kinds of HRA and HF tools do we have available and for whom? Are they tools for the specialists in these activities or are they tools for the non-specialists in these areas.

MR. SIU: If I could just add to that point. We were talking about organizational vision and I think that was something that we had shared with human factors. As we indicated earlier, PRA and human factors provide different sets of tools for different problems. Clearly, we have to address needs presented to us by the agency users. From a technical vision standpoint -- and this is where we're going to split a little bit because we have different areas of coverage, different domains. On the HRA side, if you want to talk about a very long-terms vision -- and it may not be all that long-term. I hate to think of 15 years out. Five years is kind of our current planning horizon now. I think it's reasonable to hope that we will have a common high level HRA model.

I think there's reason to believe that we can get there. When you listen to different developers talk about what they're doing, the concepts they're using are very similar. We have differences in terminology. We have some differences in scalp of particular modeling elements, but they all share very similar features.

Furthermore, I believe that there's a sense at least in a good numbers of members of the community that there is a need to drive towards some sort of common goal.

MR. POWERS: When you say "common model", you mean common with the agency or common within our nuclear community?

MR. SIU: Within the HRA community, at least the ones that perform assessments for nuclear power plants and similar facilities.

So, we would like to work towards that. That gets to George's point about knowing what others are doing. We're trying to go beyond that. We're trying to work with these others to develop this common high level model.

It's still a very high level description. You're talking about the notion of, for example, the importance of context and modeling the context explicitly. You still have to get it drilled down to what specific elements of context are you talking about. For example, are you talking about it in a static context, a dynamic context, and so forth?

My belief there is that, as now, in a few years we will still need ranges of methods and tools.

Sometimes very simple tools are good enough for the problem at hand, and sometimes you need a much more sophisticated tool. Our job would not only be to develop those tools, but also of course develop the guidance of when do you use one tool verses another?

Again, if you want to talk in terms of vision, this is I think where we might head. Obviously, there's a notion of validation involved here as well. And what Bruce talked about this morning, point us in the direction that we're going to start using -- we believe we're going to start using existing data and we can start generating new data to support at least some limited validation of these models.

I think, as I indicated in one of my answers to I think George's question, it's unlikely that we'll be able to validate these models in all performance areas. But at least for those areas where we think we can collect data, by all means, we'll try to do that.

Obviously as John Flaca indicated, we have to have a capability to address emerging issues. So the methods and tools that we're working towards now, and we have a laundry list of those, we tried to present those in that two-dimensional matrix. But we've also had a list of the issues and that appeared in that paper that we distributed before the meeting.

So those are issues we recognize that we have to deal with now, and we're trying to deal with. Certainly, things come along the path that we haven't anticipated, and we have to have the capability to address those. So, that's kind of the high level vision.

In terms of quantification in particular -- again, the HRA involves qualitative and quantitative aspects. On the quantitative side, we've been talking internally for a while about the notion of reference values, and perhaps interpolation schemes can think of it conceptually. Once we've identified what are the important factors, you define some sort of phase space, and you can hopefully through experiments or super sophisticated analysis develop some reference points to use as a basis for some sort of scheme to say what should the probability be in another part of the phase space for which you don't have those reference points.

So, that's conceptually a notion that I think we're trying to pursue. You won't see much of that in current discussions on quantification because again, we were trying to make sure that we had a good wrap-up of the expert elicitation process that we're using in ATHENA. But there are some place in, I think, the conference papers where we do take about the notion of reference points.

So, again, that's the direction of where we're heading. And I don't know if you wanted us to through the laundry list of activities that we've got to give you a sense of the breadth of applications and the particular technical challenge areas that we think we need to address.

MR. POWERS: I think your slides this morning provided a pretty good inventory on your current applications, and less of an inventory on where you think you ought to be applying HRA. For instance, we raised the issue of Option 2, if replaced, that maybe there was a rule for HRA to apply.

In some sense, I think that NRR generates user needs based on their thinking about things. I'd be equally interested in the user needs you think they should be sending to you. Do you think there's a richer field there that can be explored now, and is there yet another even richer field once you have these tools that you've been talking about?

MR. PERSENSKY: If I may, from my

perspective at least?

MR. POWERS: Sure.

MR. PERSENSKY: As far as the user needs are concerned, most of the user needs are in fact things that we as a staff talk about together. So, it's not like we're over here. In fact, we need to draw on their experience and the kinds of things that come up in the application of what tools they currently have and where those weaknesses might be.

On the other hand, similar to when I was talking about the study we did on the ROP, we indicated there that here are some things that we think might be helpful. So, it's not that we're not already doing that. It may not be to the extent that you'd like to see it, but in fact we do have that process in place and we talk a lot amongst ourselves as far as how we address that.

As Nathan had indicated, there is somewhat of a difference in what you might consider the vision between HRA and human factors though they are very related. He talk about guidance documents, and that's what we do. But I've been envisioning and I've said in the SECY that what we probably need is some sort of toolbox. With current technology, we can move a lot of this stuff that we now have on paper and pencil onto even something as simple as a palm pilot to take inspection modules with various links to be able to get into the technical basis.

So, it would be something that is useable that addresses all of the various documents that are out there right now. Human factors, as I said, is not just vanishing interface. It has all those same elements, elements that we talk about in terms of PSFs for instance or context.

So, there's that aspect of building something. The vision is trying to put everything into one place so that you don't have to carry around a bunch of paper, but also that there be an infrastructure in place that allows us to continue to develop those that need to be improved upon.

We've taken a lot of heat for 0700 in the past. Yet, it is one of the most used documents, not only by the NRC but in the industry. When it comes to control room design, the EPRI meetings, most of what they're doing in developing their stuff is based on that. Nonetheless, it could be a more useable, more useful kind of document. There are still gaps in it. There are still things that we don't have good guidance from.

Most of the guidance that was developed or put into that document came from things that we stole from the military. This is not that we did a lot of research, in terms original research in a laboratory to develop that guidance. Most of that guidance was taken from other places, but we went through a validation process.

The few things that we were able to do in a laboratory type setting, we've made use of the Halden project and whatever we could to get simulator data and develop the guidance and the criteria that are established in those documents. So the infrastructure is really something that -- whether it's our simulator or Halden's simulator or some other simulator, we need access to that kind of thing for operations.

The thing that we have somewhat ignored by spending a lot of time on simulators is that a lot of the errors, and one of things that we found in some other studies that we did with INEEL, was the issue of latent errors. Those errors were being made by maintenance people, not by the operators.

MR. ROSEN: That's my opening.

(Laughter.)

MR. ROSEN: In the context of tools, what you spent more of your time on, and I think appropriately, is the focus on control room operator performance. But what Davis-Besse tells us, and what a lot of other stuff tells us, is that personnel outside the control room, including top managers, maintenance people, supervisors, and engineers can make mistakes too. Mistakes they make become latent errors, and those are the cases that come out and bite your leg.

So the question here, in the context of tools, what tools do you need to look at the performance of other people who are not control room operators? And this gets to the question of organizational performances or rich literature, which I'm sure you know better than me. There's rich literature on organizational development in psychology and how that factors into the personnel performance of engineers and managers and all kinds of people in the organizational settings, and what sort of tools should we be using.

It think that this is the opening. This is the area that can have the single biggest incremental value to the agency. I know it's controversial. If it wasn't controversial, we probably wouldn't be interested in it.

MR. APOSTOLAKIS: If you do it in the context of how to these things affect human, I don't even think you need to go to the Commission.

MR. ROSEN: Well, that's what I'm doing about. Organizational performance is safety culture. And organizational performance is simply the sum over the integral of all the individual performance.

MR. APOSTOLAKIS: You're doing it because you're trying to understand human performance. There would be no objection. That's the way I understand it. I'm serious.

MR. HALLBERT: Part of the --

MR. APOSTOLAKIS: But if you say, I want to establish a program of safety culture, you might as well not even call. You shouldn't start it by itself. You should start it in the context of something that is immediately useful to the agency.

Yes, Bruce. I'm sorry.

MR. HALLBERT: That's okay. Part of the insights from that work that we performed on the errors in power plants that contributed to these risk significant events did identify that maintenance errors were important contributors to many events.

One of the questions that we entertained when we were back here at a meeting on that particular project was if we could just eliminate maintenance errors, could we make a substantial improvement in reducing the number of risk significant events. In other words, if you needed all the failures that occurred in this event for this event to have occurred, if you just removed maintenance errors, you would thereby reduce the number of total events that had occurred.

Part of the quandary in an approach like that is recognized in that maintenance failures for maintenance contributions to significant events don't occur in a vacuum of maintenance. They occur in a context of the overall plant division of responsibilties and mission activities. They're linked to engineering activities, they're linked to operations activities, and it seems like -- and this is just maybe just my opinion right now that I'm saying -- but it seems like if you want to get reductions in the overall rates of some of these kinds of events, you have to understand those contexts and go into some of the causes of those maintenance errors, just like the kinds of causes that contribute to corrective action program failures.

MR. APOSTOLAKIS: They're not just maintenance errors.

MR. HALLBERT: True.

MR. ROSEN: The reason that David-Besse didn't find the problem was because there was an error repeated several times in putting in the access ports. That was an engineering or a management error. If they had put the access ports in, then maintainers would have gone and said that stuff is coming from something other than the flanges.

MR. APOSTOLAKIS: As I said earlier I started reading this root-cause analysis, which is very good. To make it interesting, I started making notes.

If this deficiency can be identified, what is it telling us? Some of them are telling us that the work processes were not very good. They were not required to do certain analyses after they found something, you know. That's a relative easy fix.

I think where the main difficulty will be when they know of the problem and they don't take action. Because, I don't know how to model them. I think that's going to be more difficult. They say it very clear, "the plant restarted without taking correction action for identified problems." This is the utility speaking now.

But these are the kinds of insights that are beginning to address these questions of organizational questions and so on. I believe that as a community we spend too much time trying to model errors during accidents sequences. It turns out that pre-initiating events are much more threatening.

MR. ROSEN: Well, I don't exactly agree with that. I think we spent an adequate amount of time on operating sequence. But, we spent almost nothing on the other piece. I could not do what we've done. We had to do that. But we spent almost nothing on looking at errors.

MR. APOSTOLAKIS: When people talk about errors of commission, automatically they think of a sequence or something that's happened already.

MR. SIU: Just as a comment here -- actually, this is one nice case where feedback from the human factors work led to a task in HRA. We have a task on latent errors, which doesn't get to your point George about the cause and initiating events, but the notion there was to start exploring again the issue of latent errors.

There were some beliefs -- in fact we talked about this issue in Stockholm back in '95 or something like that -- that we have at least HRA tools to deal with the likelihood of, for example, maintenance errors. There was some feeling that the THERP methodology was just fine for that kind of application. Now that was stated without any strong technical analysis but it seemed to be reasonable to the people attending.

What wasn't covered there was the notion of the dependants between multiple errors. Now you start asking about the underlying causes, whether it's culture, whether it's work processes. We intended to look at work processes as part of this work.

We haven't gone as far as safety culture. But now that we heard from Scott this morning, we'll probably open that up and see if we should approach the Commission on that.

MR. APOSTOLAKIS: Again, I don't think it would be wise to say we want to study safety cultures.

MR. SIU: Right, but as a contributor to --

MR. APOSTOLAKIS: Right. We are doing this, we have started it, and now we have to move into this area. You know, that kind of thing.

MS. LOIS: I just want to mention although it's in a past life, the University of Minnesota had done some work in the early 90s, and the early indications were that learning and management commitment were very good predictive indicators.

MR. APOSTOLAKIS: No, I understand that. But I really think you have to test these things against what they found in Davis-Besse.

As I say, some of it is just "all I have to do is fix the process". Some other things though, the knew of problems and didn't take action --

MR. ROSEN: Well, there's a corollary here, George. Just looking at Davis-Besse is not enough. One needs to take some hypothesis out of the Davis-Besse circumstance and then apply elsewhere. And one of the place was Indian Point.

If you think about Davis-Besse, they didn't put the access ports in and they could've. Now Indian Point didn't replace the steam generators when they could have. And so again, you come to the question that there's some commonality.

MR. APOSTOLAKIS: Absolutely. I just mentioned Davis-Besse because it's a hot issue, and I just happened to get the root-cause analysis a week or so ago and I was going through it.

But even there, you say your talking about the access ports, that they didn't do it. Maybe they didn't do it for a long time. They were deferring it from outage to outage for three, or four, or five times or whatever. Is that the indicator or the fact that they didn't do it at all?

These are the kinds of questions that I think the researchers will have to answer. Some people are saying a good indicator of a safety culture -- not the total of course, but a good indicator -- is the number of items deferred. They were planned to do it and they were not done during the outage. So, there may be ways to approach it and get some indication.

MR. POWERS: Let me see if I can summarize what we've said about tools.

We have not a great deal of schism between HRA and HF here, but some. That in the HRA, you're looking to develop tools of varying levels of sophistication and the guidance for selection among those tools, that you're looking to validate these tools both by using existing data and Dr. Bonaca has suggested that we look to see if we can use the data for development of symptomatic procedures.

I'm less persuaded that we will have access to that data or even that this data is readable to this point. It seems to be a common problem when getting the data collected over a decade ago that it is no longer readable by any machine that we have. In some respects, what we may be discovering is that we've gotten sophisticated enough that the controls in that data were too loose to make it very useful to us. So I'm less enthused about that, but it's worth looking for.

But more importantly, you're looking at can we develop data to develop new data to provide some sense of validation recognizing that validating these tools that they use strictly in an interpolative fashion is a pipe dream and it's never going to happen. You may be able to find some reference points in a space that you have some confidence in, and you're hopefully no extrapolating vast distances.

Now what we learned just before lunch, that phase space you will of has dimensions that perhaps we haven't explored yet. We don't know what they are because we have variants in the data and you can look upon variant data as projections from the space that has a high dimensionality.

In the HF area, we're looking at a somewhat different kind of tool, more user-oriented, more delivered to the frontline kind of tool that's the implementation of a vast amount of technology that's in hand now. Is that my understanding?

MR. PERSENSKY: That's part of it, yes.

MR. POWERS: You go on and say you want an infrastructure that allows you to build upon that, but the tool that you're producing is one that would be used not by a specialist but by a non-specialist.

MR. PERSENSKY: In the end, yes, that tool, as well as the training that would go with it.

MR. POWERS: And the training. You still have a guidance aspect to this?

MR. PERSENSKY: Right.

MR. SIU: If I could just add to what you said, Dana. Again, it's not that we're not going to also develop guidance for non-HRA analysts. Again, someone who's reviewing an application wants to know from an HRA perspective, so we're also trying to address the user.

MR. POWER: Yes well, that point that you made, that I took a lot of notes on that I don't see right now, we are tying to support NRR, who are doing the -- I really put that under your guidance category rather the tools category.

MR. SIU: Okay, parse anyway. But there's one thing that says here's guidance, how to use this set of HRA tools. Here's the guidance which might support or review of somebody else's --

MR. POWERS: That's right, and I made a

distinction between the two.

MR. SIU: Okay.

MR. POWERS: Now a question that was raised in connection with your data, do we need our homegrown simulator? You know, the simulator for a US plant run by US people doing the kinds of experiments now done by a Norwegian simulation of a Finnish plant with Finnish plant and Swedish scientists.

But the question posed to you is: without thinking of cost benefit right now, could the research program make bigger use of that kind of a facility?

MR. APOSTOLAKIS: I'm a bit confused. How is this facility different from the simulators that exist right now in this country?

MR. POWERS: This is a research simulator. They go do these wonderful tests and things like that. They invite crews to come spend a wonderful week in Chattanooga running experiments for them, wired up like Ginny pigs with stress measures and stuff like that. I mean to develop data, to develop an understanding, to develop a science.

MR. PERSENSKY: The issue is it's a reconfigurable simulator that you can change things around, which you can't very well do at existing plants or even at our own simulators in Chattanooga.

In addition, there would be much wider data collection opportunity, the kinds of things that they do have at Halden and other facilities, NASA facility, FAA facilities. They'd collect a tremendous amount of data. We never even talked about the data they'd collect. That might get into much more finite kinds of things.

But to answer your question in the best of all possible worlds, having a simulator like that I think would be helpful to human factors. It would be helpful to HRA and it would be helpful to Digital I&C at least. I don't know really that it's that practical.

MR. POWERS: The answer is unequivocally "yes" to the question that's posed. But the follow up question is: do you have a strategy that would make use of this, and would it make use of it 60 percent of the time, 70 percent of the time, 100 percent of time?

MR. APOSTOLAKIS: You are asking uncharacteristically an unrealistic question there. I can't believe my ears.

MR. ROSEN: It's not his question. It's mine.

MR. APOSTOLAKIS: Divorce always from cost. Maybe it's cheaper to fly US troops -

(Laughter.)

MR. APOSTOLAKIS: I mean you're asking would it be nice to have this extra research capability. I'm not going to say "no". It would be nice.

MR. HALLBERT: I guess, you know, from another research perspective also, it depends upon the kinds of questions you want answers to.

For example, you talked earlier about the data available from EOP studies for relicensing and requalification exams. If part of what you want to do is collect a larger baseline on operator performance in different contexts, there probably is a large amount of suitable data there.

If what you want to do is something more unique that requires modification of the operating environment, then you have to start looking at the extent of modifications and finding out can it be accommodated in the existing facility.

If, for example, what we were talking about doing -- and I'll use an example here - evaluating how well a new electronic procedures system would work. Well, you wouldn't actually have to have your own dedicated plant to do that because a number of plants considering doing that right now. You might try to find a plant that was interested in that and say I've got a couple of candidate systems we want to research, can we use your training facility.

MR. ROSEN: The answer would probably be "no" because it's used 24 hours a day, 7 days a week. You kept coming back to your own point that those simulators are heavily used.

And licensees, it's crucial that they get the training done that they have scheduled. They can't afford to have somebody in there messing around with their simulator because at seven o'clock in the morning, their crews are coming in.

MR. HALLBERT: So you'd like to piggyback on efforts that are already going to try to take advantage of data that they're already generating. But unfortunately, the problem that we've always had in the past was something like this, that it is not a regulatory issue.

Very few plants want -- well, I'm not sure how many or which plants like to volunteer for that because if something happens during the simulator exercises that they don't like, then it immediately raises issues for them.

MR. POWERS: And you're never going to find a plant that has an appropriate simulator for looking at a modular plant.

MR. ROSEN: But the point is if we don't ask these questions, if we don't ask them now, they will not be asked. Here we are at the verge of perhaps a new generation of reactors, we all hope -- are we just going to do it the same way we did the last generation, or are we going to do it a little differently?

MR. POWERS: Well, I'm kind of impressed with the last generation lately.

MR. ROSEN: I think we ought to do more. It look 50 years to get to the point where the old generation -- it's pressingly talked about.

MR. POWERS: And now you want to put in another new generation to get me depressed again. You're playing with my sanity here.

MR. APOSTOLAKIS: This is not the only way.

(Laughter.)

MR. ROSEN: To start off, this generation of machines, if we're going to build advanced reactors, highly integrated control rooms, passive safety, it seems to me that an investment upfront of what it takes to build a reconfigurable machine where we can test some ideas and test these things is not entirely out of the question. It shouldn't be.

MR. POWERS: If we're going to ipso facto attack the issue of errors of commission, I don't know how you do it if you don't go get some exploratory data. I mean everybody just throws up their hands for error of commission, and I think exploratory studies may be the only way to broach that subject.

MR. PERSENSKY: If I may, one of the efforts that I put into the advanced reactor plan, the first effort in that included sort of a scooping study of what might be the problems with advanced reactors that we should be addressing, where the gaps between what we know, what guidance we have available, and where we might be going if there's a need to chance. For instance, for advanced light water reactors, we may not need to make many changes to the current guidance. For modular reactors, we might.

But in that, we included an element of looking at the need for a simulator. One of the things that we talked about in that particular element of the plan was that currently we've got "X" plants or units out there. Each plant has its own plant specific simulator, but they're all different; whereas for the future plants, we're looking at more commonalities.

So, there might be a real possibility of joining with the industry or with DOE in developing a simulator that we can all use. Not unlike the kinds of things that they did with some of the test facilities with some of the vendors, where we were jointly funding and working towards that.

So we are interested in that, and we plan to look at that as a matter of fact.

MR. BONACA: But I think you want to have a simulator of a plant with a matching set of procedures for that plant. If you build a new simulator that maybe wonderful as a concept -- but you don't have the procedures which are tied to the machine.

One suggestion. A number of plants have been retired, but they had plant specific simulators. They're probably still effective and can be used.

MR. PERSENSKY: They've all been bought up or trashed.

MR. BONACA: Okay.

MR. PERSENSKY: Because we purchased a couple of them for the TTC as a matter of fact. Some of the others had been purchased by other vendors.

MR. GRIPMAN: I'm Dave Gripman. I wanted to comment on -- Jay stole my thunder there, but I think this idea of looking for synergy with the Department of Energy is a petition to cite a pebble-bed reactor. They have a lot of operations experience and operators available.

I think that might be a way to do some cost sharing because I think the use of this research simulator is a very powerful one. I think having one in the US in addition to whatever else we can learn around the world is a good concept. We can full scope. We can look at test simulators and extract general principles and behavioral profiles as well for crew performance. So, I think that's one way we want to go.

I think the other challenge has to do with the issue that was raised a little earlier on maintenance. When we talk about a simulator, I think if we're talking about simulation, we almost have to go to analytic type simulation if we want to talk about maintenance performance, looking at work processes, and what happens when you disrupt time. Can you force common cause failures across systems and look at what those failure rates might be like to see if those shaping factors were the same?

That's a more challenging type of simulation I think, and that's something that maybe ought to be pursued as well.

MR. POWERS: Peter?

MR. FORD: The answer to my question has been partly answered at least to advanced reactors. If we believe the schedules we're seeing, within the next two or three years, we'll be looking at applications for designing new reactors. We don't have simulator for these new concepts. Therefore, you have to rely on the synergy between the conventional reactors and the new reactors that are coming down the line.

When you look at your needs over the next two years, what's keeping you awake at night? You have no way of knowing how you're going to tackle a particular problem in both the human factors and HRA. What keeps you awake, the sufficient lack of knowledge?

MR. SIU: You know what keeps me awake at night? Nine-eleven.

MR. POWERS: I'm jumping to speak here, which is silly on my part, but I do rather silly things. But when I see massively automated plans, I put on an HRA or a human factor hat, and it's the errors of commission.

I probably should probably worry about

latent errors in the maintenance process. The committee definitely heard the story that they were four times as important as the errors following an initiating event. We got that message last year and we quote it frequently.

MR. APOSTOLAKIS: I'm a little surprised though that some committee members seem to be more enthusiastic about getting the simulator. Rosen and Powers are saying this is great.

MR. FORD: Hold on, George, before you get into that particular topic. Nothing keeps you awake at night?

MR. APOSTOLAKIS: I'm not going to say it now.

(Laughter.)

MR. HALLBERT: I not sure it keeps me awake at night, but it's in my thoughts in the daytime when we think about HRA and we're going this work. I have children so they keep me awake at night.

(Laughter.)

MR. POWERS: Wait until they become teenagers.

MR. HALLBERT: We have that too. They wake us up at night when they come in.

(Laughter.)

MR. HALLBERT: Just a couple of things. One is just trying to reconcile the notion of reliability and validity in the approaches that we currently use. I'll give you some examples.

Reliability is different analysts being able to replicate the results, looking at the same scenarios with the same information. There have been some benchmark studies in which the orders of magnitude difference in results is really bothersome. You know, where they did try to benchmark.

MR. POWERS: There is a really nice paper which I had read, but I cannot refine, in which they compared some of these analytic techniques to each other, and it -- human reliability analysis, and it virtually --

MR. APOSTOLAKIS: It was all over the place.

MR. POWERS: Yes. I mean there was no correlation whatsoever.

MR. HALLBERT: The other thing is just the validity for -- I'm not sure if I'm characterizing this correctly, but at least to me, an apparent lack of a process in which methods become validated. In other words, a group of people produce a method and it's then just released.

I'll say this for ATHENA, ATHENA at least has gone through a lot of very systematic attempts and efforts to try to achieve some kind of validation of the principles of the method. Just given all the methods that are out there, there are some methods that have done that to a much less extent, so you really wonder about different analysts using it. You wonder about the validity of the results that come about as a result.

I then think about the NUREG on lessons learned from the IPEs. And in the appendix, I think there's a very -- I think in fact you wrote it Dana if I'm not mistaken or at least you talked about it at the EHPG in Norway I think when you came over there. There are certain criteria to a PRA completeness. And with regard to HRA, there should be the same criteria. So, I don't think we're there with HRA yet.

MR. APOSTOLAKIS: The thing that really bothers me, and it comes to my comment earlier, is that, as I said earlier, I read one model and they seem to be focusing on decision analysis. Another model is focusing on time. Another model, it says PSS. Another one is expert opinion. And, they operate in parallel with apparent interaction. I think it's time to stop that.

MR. FORD: So follow up then, on both the HF and the HRA, you've got data collection analysis. And you're saying that keeps you awake as you go forward on the current fleet using it in its entirety going to advanced fleet. In the prioritization of tasks for the next five years, is that item high on the prioritization list, data collection?

MR. SIU: Practically number one.

MR. FORD: I haven't seen it yet, so --

MR. POWERS: Nathan says it's number one on their list.

MR. FORD: Great.

MR. SIU: That and guidance are the two tasks that we are really focusing on.

MR. POWERS: To follow up on George's point, my understanding of your program is that you know have, you have number one, guidance. Number two is this data collection. Somewhere down a little lower is to look at all these models, distill which are the good aspects, which are the bad aspects, and come up with some judgment on what a desirable tool would be. Now that may be one that already exists, or may be one that you have to invent, or it may be that you can change a Greek thing into a Latin thing.

(Laughter.)

MR. SIU: He says that, but it's a step backwards.

MR. POWER: Okay, a Greek thing into an Anglo-Saxon thing, which is clearly a step forward, and have a new model.

Is my understanding correct there?

MR. SIU: Again, I think we're talking about, as you indicated earlier, is a range of methods and tools suitable for different applications and guidance to support the appropriate application of those methods and tools.

George, I don't think you were in when we were having a little bit of discussion about driving towards some sort of common model. That's something I think that we would really like to do.

MR. APOSTOLAKIS: Good.

MR. SIU: Some of the discussions we're going to have next week are along those lines.

MR. APOSTOLAKIS: Very good.

MR. POWERS: I very much appreciated your presentation. The information was enlightening to us and extraordinarily useful. I wish you well on whatever follow-on efforts you're taking.

MR. APOSTOLAKIS: Keep your passion burning.

MR. KRESS: And get some sleep.

(Laughter.)

MR. POWERS: I appreciate Nathan sharing that material with us because it was helpful on many, many scores.

MR. SIU: While Bruce and Dave are packing up here, another thing I wanted to mention by the way, you had asked about, if you will, the gaps in our program.

MR. POWERS: Yes.

MR. SIU: What you see in Erasmia's slide I think are, most of those are anticipatory activities. For example, the latent errors, we talk about extended applications for LOPAR, and shut down long-term recovery actions, level two HRA. These are things that we are anticipating that we're going to need to improve methods and tools for. Obviously, we've got stuff being used now. But the question is can we do better.

So the list you see in the table that was displayed is our shot at what we think the needs are. We have something that's very global on upgrade and advanced reactors. Maybe it's not specifically enough --

MR. APOSTOLAKIS: On page 19 of the

plan, you have a number of tasks.

MR. SIU: That's right.

MR. APOSTOLAKIS: These are the same?

MR. SIU: Those are the same. We just tried to map those into different needs.

MR. APOSTOLAKIS: Okay.

MS. LOIS: Except, a few tasks are not there such as standards development, vulnerability, or --

MR. SIU: That's right. So, there are a couple of things that have been added on the table.

MR. APOSTOLAKIS: There is also some acronyms at the end WSMS 1-2.

MR. SIU: Yes.

MR. APOSTOLAKIS: RSWER 1-3. Is this a secret code?

MR. SIU: No, this is our risk informed regulatory --

MR. APOSTOLAKIS: That's the RIRIP. I understand that.

MR. SIU: Okay. And it has specific activities in it, so these are teed to those activities. So when there are activities that need HRA support --

MR. APOSTOLAKIS: I have two questions

regarding this table, appreciating the fact that it's in a document dated May, 2001.

MR. SIU: Yes.

MR. APOSTOLAKIS: One is, would Davis-Besse or the Indian Point incidents, among others perhaps, change these tasks because that was done under a different context?

And second, I understand you plan to have an updated version early next year. I think that developing performance indicators for human performance is important. Maybe you can try to accommodate this somewhere there because the reactor oversight process is in desperate need of this. It does relate of course to Davis-Besse and Indian Point again.

Again, I don't mean performance indicators in the sense that they are already in the ROP for reactor safety like the frequency of transients of the frequency of this and that because you may not be dealing with frequencies.

But when the guy there to inspect, is there an indicator that he can look at? Like I mentioned, a number of items deferred for example. Does it make sense universally? But I really think these are what the issues are these days. So other than that, it seems to be a fairly comprehensive list of various tasks and theories.

And one last comment I keep forgetting. Jim Riesen I think makes a distinction between latent errors and latent conditions. I think latent conditions is probably more appropriate because they're not necessarily errors. They create the context within which -- it's a broader term. I think conditions is a little better.

I have a few other comments on the report, but the report seems to be obsolete anyways. For example, on page 20, there are some deadlines.

MR. SIU: Yes.

MR. APOSTOLAKIS: "Develop HRA research lessons to support risk informed regulatory applications", September, 2001. Has that been done?

(No response.)

MR. APOSTOLAKIS: "Develop initial guidance" -- well, there are certain things that are supposed to be done by now.

MR. SIU: Right.

MR. APOSTOLAKIS: And I wonder whether they have been done and if we could get copies of them.

MR. SIU: And as Erasmia indicated, the

two things that are coming down in terms of quantification uncertainty, and that's in the context of ATHENA, and what was the other one? Oh, PTS was the other one.

But yes, the plan will be updated. Obviously, one of the motivating factors behind that is because the dates need to be updated.

MR. ROSEN: When Scott came at the very beginning, he tantalized us by saying we may need to reengage the Commission on Davis-Besse, based on the Davis-Besse experience. Is there more that you can say about that? Is there a whole piece of this presentation that hasn't been given or what?

We have said a lot about it. George has spoken, I have spoken, and people have said things around the table, but you haven't said anything.

MR. SIU: We haven't done significant work in the area. The decision that we would think about reengaging was a very, very recent decision. This is a statement of intention I think, and we're going to start looking at that.

MR. ROSEN: Perhaps you might need some input, more than we've given you already.

MR. SIU: Sure, yes.

MR. ROSEN: One of the pieces of input I've mentioned before was the leading indicator program at EPRI. And the offer that the EPRI management made to me at least was that they would be pleased to come here and brief the staff and the ACRS if we wanted to and the subcommittee on what that program does.

To me, in looking at it and talking to one of the leading utilities that's using it, it's the first piece of data collection that in mind the industry has done that actually has a chance of getting us an early signal that the decision-making environment in a utility is degrading, that tasks are not being done well. I think that's a piece of this problem, an organizational performance problem, that we're labeling safety culture.

The other thing is we talked about the need for indicators. Well, even leading gives you these indicators, to sum it up and look at things. But George mentioned the modifications that are preferred. To me, just corrective actions that are preferred that are significant is another one of those indicators that are important.

Of course, the classic one in corrective actions is the failure to preclude recurrence. The very essence of a corrective action program is that when something happens, you do enough to make sure it doesn't happen again. And when it does happen again, if it does, there ought to be a big signal to management that something is wrong with the corrective action program.

And the third one is a question of, in an environment that is degrading, in a place where there are a lot of good people, those people begin to come forward. In a safety conscious work environment, those people come forward with complaints that we're not doing a good job. How many there are and what management does with them is another indicator of the degrading environment or an improving one.

So, there are some rich data sources to mine. To me, working on how good the operators do in a known transient -- and it's a good thing to do, but it's working on a problem that we've worked, and worked, and worked. We haven't worked at all hardly on this other end of the real risk spectrum.

MR. POWERS: I expected you to -- I mean you certainly mentioned this leading indicator program and its value. But I expected you to go on and comment on this whole business of cross-cutting issue, and how is the HF and HRA program addressing this?

I mean you've got this statement. This is a cross-cutting, and it just kind of sits there.

What do we do with that? I mean is there nothing that can be done?

MR. APOSTOLAKIS: In fact, there was a hypothesis which the ACRS several times in its letters said it's an untested hypothesis. That is there is a problem with any one of these three cross-cutting issues, we will see it in the performance of the hardware so why worry about it.

MR. ROSEN: To my view, that is exactly correct. If there is a problem with cross-cutting issues, you will see it in the hardware. The trouble with that is that you will see it too late.

MR. APOSTOLAKIS: Too late --

MR. SIEBER: The other problem with that is you're not going to find just one issue. You're going to have a whole series of latent defects in the plant that will take you millions of dollars to correct and years to correct.

MR. ROSEN: And the other point that you will apply but didn't make is that if you have a whole raft of these defects, on a bad day they'll all line up wrong. Then, you can have a very serious circumstance.

MR. APOSTOLAKIS: Like Swiss cheese.

MR. ROSEN: The barriers all have holes,

and then one day the barriers all line up exactly right and you get this light --

MR. APOSTOLAKIS: When they say "model", that's what they mean.

MR. POWERS: The ROP people, when they respond to us -- and this is untested hypothesis -- said "yes, we're going to test it", I don't know how they can test it without you people being involved.

MR. PERSENSKY: To some extent, the report that I mentioned that talked about the ROP study, which is NUREG CR-6775, was a response to that question. They did look at how performance was characterized in the reactor oversight process and how it lined up ASP events in the past. That did identify a number of issues.

The one that seems to have the highest payoff right now is the improvement to the corrective action program inspection module. What we're doing is looking at the inspection module.

It did mention some other issues that came up. For instance in the area of latent errors, the possibility of some changes to the sampling under the maintenance program, the maintenance rule. There are certain things like we look only at certain high-risk equipment. Whereas if you look back at some of the accidents, there were other pieces of equipment, that when they lined up properly, caused the problem. So there may be some other changes. We proposed that we look at that.

Also, the issue of communications is one of things that came out as a major problem. But we do have in fact right now, since that work was done, we have come out with a couple of reports in conjunction with NRR on trying to improve the communications' look at things. So, we didn't go back on that.

We also mentioned what might be called safety culture. We made the point in our letter that there is a current restriction on doing much work in that area. But as Nathan said, there's very recent direction that we may be going back and looking at that.

So, there are a number of things that came out. If you look at the three cross-cutting issues -- one is the corrective action program, one is human performance, and the other is safety conscious work environment -- they're all human factors.

MR. POWERS: They're all one thing.

MR. PERSENSKY: They're all one thing. They all come down to a human or organizational or whatever factor.

But, we have done some work in that area. We haven't done perhaps the definitive work, and I think we need to follow it up with more recent looks at things like Davis-Besse.

MR. APOSTOLAKIS: I have a question to that regarding the plan. There was a conceptual problem I had with this.

It says that the methods for modeling or post-initiate actions are in not fairly good shape, but they are more advanced than methods to treat organizational factors. Now we all agree that organizational factors, as the report says, strongly affect those actions.

So how can a method or action be more advanced than methods for dealing with something that's necessary to understand the actions themselves? If I do organizational factors poorly, don't I automatically do human error modeling for which organizational factors are important?

MR. SIU: Or put it another way. Perhaps you're dealing with some sort of an average level. I mean you're able to distinguish between the characteristics of different organizations other than how they affect things that we do try to address in the analysis. Like when we make observations of crews and see how they actually respond to a particular event, or you look at past history and factor that into your analysis. But that's not a direct analysis of --

MR. POWERS: I think you see it in a great deal of variance in the data that you collect on human performance. If you don't understand everything and you project it under the space that you understand, you're going to see a large amount of error. And that's what they see.

MR. ROSEN: They do not understand what the source of the variance is.

MR. POWERS: That's right.

MR. APOSTOLAKIS: What I think really is said here is that there has been a lot of attention paid to modeling human actions. There are a number of models. In that sense, it's more advanced than the other stuff where you have maybe a couple of models. But, it's causing effect. If the cause is not modeled well, the effect is not modeled well. But again, I do bring it very serious.

I have a question for the Chairman.

MR. POWERS: Yes.

MR. APOSTOLAKIS: What time does the coffee shop downstairs close?

MR. POWERS: I believe you will not be able to get coffee after four o'clock.

MR. APOSTOLAKIS: Okay.

MR. POWERS: Let me ask this question. I had five categories of questions that we posed after lunch: the big scheme of needs, tools, organization safety culture, and indictors, development of HRA models and view of existing models, and state of the art. I think we have addressed those in our discussions.

Do you want to take a break for 15 minutes, get your coffee, come back, and do a roundtable for the points that we want to make?

MR. APOSTOLAKIS: Sure. I think that's good.

MR. POWERS: Or do you want to interrogate these gentlemen and lady further?

MR. APOSTOLAKIS: No, but I'm sure they're going to stay.

MR. POWERS: They're more than welcome to stay because I think we're going to need their continuing help.

But I will emphasis that on the time that I have been on the ACRS, this has been the most enjoyable, pleasant, and well thought out meeting in the area of human reliability and human factors that I've ever attended. It comes off with a more optimistic note than I've ever enjoyed.

So, I congratulate you on an excellent presentation to the subcommittee, which almost amounts to the full committee. You will be surprised to find that Dr. Shack, who is not here, has strong views on this subject and will probably take an orthogonal view on everything.

We do need to chat a little bit about what to present to the full committee.

We're done. I think at this point I'm going to close the meeting, and adjourn this transcriber at this point. We'll come back after coffee and discuss a little bit about what to present to the full committee and what we think ought to appear in the letter. So why don't we reassemble at twenty-five of the hour.

The meeting is closed.

(Whereupon, the above-entitled meeting concluded at 3:19 p.m.)