# Official Transcript of Proceedings

# NUCLEAR REGULATORY COMMISSION

Title: Briefing on Human Reliability Program Activities and Analyses: Public Meeting

Docket Number: (n/a)

Location: Rockville, Maryland

Date: Thursday, May 29, 2014

Work Order No.: NRC-814 Pages 1-159

**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

+ + + + +

BRIEFING ON HUMAN RELIABILITY PROGRAM

ACTIVITIES AND ANALYSES

+ + + + +

PUBLIC MEETING

+ + + + +

THURSDAY

MAY 29, 2014

+ + + + +

The Commission met in the Commissioners' Conference Room, 1st Floor, One White Flint North, Rockville, Maryland, at 9:00 a.m., Allison M. Macfarlane, Chairman, presiding.

PRESENT:

ALLISON M. MACFARLANE, Chairman

GEORGE APOSTOLAKIS, Commissioner

WILLIAM D. MAGWOOD, IV, Commissioner

WILLIAM C. OSTENDORFF, Commissioner

KRISTINE L. SVINICKI, Commissioner

ALSO PRESENT:

ROCHELLE BAVOL, SECY

MARGARET M. DOANE, OGC

RICH CORREIA, RES

EDWIN S. LYMAN, UCS

SEAN PETERS, RES

MARY R. PRESLEY, EPRI

CLAIRE TAYLOR, HRP

JAMES VAUGHN, Nine Mile Point

MIKE WEBER, DEDMRT

SUNIL WEERAKKODY, NRR

JOHN WREATHALL, John Wreathall & Co., Inc.

A G E N D A

P R O C E E D I N G S

9:03 a.m.

External Panel

CHAIRMAN MACFARLANE: Good Morning. Hope everybody's good today. I'd like to welcome staff, industry, members of the public who are here for today's meeting on Human Reliability Analysis. That's what we're going to be focusing on.

The NRC has been moving to increase the use of risks insights in our regulatory framework, and central to this effort has been use of probabilistic risk assessments to drive quantitative measures of risk, and among the items assessed in event sequences is the reliability of operator actions.

So given the increasing influence of PRA in the NRC's regulatory processes, I believe it's important to fully understand the state of human reliability analysis and the uncertainties associated with this analysis.

So today we're going to have the opportunity to look at the field of human reliability analysis in general, and to discuss efforts to develop the integrated decision tree human event analysis system methodology. So today the Commission's going to be briefed by two panels, an external panel and an

internal panel, NRC panel.

So first in the external panel, we're going to hear from Mr. John Wreathall, president of John Wreathall and Company; Dr. Claire Taylor, who is the Senior Scientist at the Halden Reactor Project; Ms. Mary Presley, the Project Manager/Technical Leader of the Risk and Safety Management at the Electric Power Research Institute; Mr. James Vaughn, the Operations Shift Manager at Nine Mile Point nuclear power plant; and Dr. Ed Lyman, who is a Senior Scientist, Global Security Program at the Union of Concerned Scientists.

So I look forward to the presentations of the panels. First, let me see if any of my colleagues have any opening statements.

COMMISSIONER MAGWOOD: Just quickly, Chairman. We had scheduled this briefing some months ago and it was cancelled due to inclement weather, as I recall, and both Dr. Taylor and Mary Presley both came in. Of course, one came overseas and one came from the across the country, and I appreciate that they are back here again today.

Several of us did have an opportunity to sit down with you when you were here before. So thank you again for returning and making the special effort. We really appreciate that. Thank you, Chairman.

CHAIRMAN MACFARLANE: Anybody else?

COMMISSIONER APOSTOLAKIS: Well yeah. It's just that impressive how popular the subject is. It's popular with us.

CHAIRMAN MACFARLANE: Okay. Alright. Well, on that note, we'll start off with Mr. John Wreathall.

Current State of HRA Research

MR. WREATHALL: Thank you Madam Chairman, Commissioners. It's a pleasure to be here. I did send in a summary of my history. But there was a couple of things, given the sort of change in emphasis from the original meeting that I wanted to mention, that my background and academic training is in engineering, not in human factors.

So I come to this with degrees in Nuclear Engineering and Systems Engineering, rather than the field of psychology, even though that's the sandbox I tend to play in quite a bit. As such, I have worked in nuclear power plants in the UK, doing hand fuel loading, all sorts of hands-on things in the plants before I moved into the consulting world. So I do have some body of knowledge and experience hands-on.

So if I can maybe start going through the slides. I have three or four topics in general and

perhaps we'll give a little more emphasis, a little less emphasis, given the members of this panel, who will cover some of the same things.

On the next slide, I'm just highlighting that right now, there is very limited development of new HRA methods. In fact, this agency is probably the leader right now in the development of HRA tools and methods, not just for nuclear power plant operators in the normal Level 1 PRA mode.

The IDHEAS method that's going to be presented later and an associated method that I think is referred to as the generic HRA method, are being developed by your staff. There is the fire HRA guidelines work.

There is the work going on to develop methods for the Level 2/Level 3 PRA, and right now there is a new reg in development that discusses human error and human reliability in the field of the medical applications.

I think that's not had a lot of visibility, but it's an area that's yet another branch of HRA being developed within the agency.

As far as overseas is concerned, there are new methods being developed in South Korea related to the use of computerized control rooms, and the next

slide, Slide 4. The French, Electricité de France has been building on its earlier work in the MERMOS technique, which is based on operator simulator trials, gaining insights from that, and building HRA methods, is being added to both for the new types of plants, which are not included in the current simulator spectrum, methods for designing as well as the PRA application.

So pre-accident human error, HRA to optimize design, activities in the design phase and also Level 2, fire PRA, seismic and so on. So the French are doing a fairly large amount of effort too. But those are the main activities and new methods.

What has been going on, Slide 5, is two fairly large reviews within the HRA and PRA communities, of methods that are already developed. The UK, as then was HSE, identified over 50 methods in use back in 2009, and the number has increased. So I see it is a time when there's a rationalization and refinement of methods, rather than further new methods being developed.

These two reviews, contributing to that, to give where the strengths, where the weaknesses are, how they might fit together in different ways, and particularly the Nordic/German/Swiss evaluation, the exam HRA is particularly aimed at putting together a set of methods that are particularly focused on the PRA

knowledge, insight and use, not just in creating numbers.

There are some backup slides on that if it's an area of interest, but I don't intend to say more than this right now in the slides, the front slides. I think given the interest that's been expressed to this panel about the development of the IDHEAS methods and the letter that was written by the ACRS, which has been sent to us, I wanted to try and clarify what I see as a discussion going on that I think is an underlying issue.

Slide 8 is the introduction to this. That we talk in HRA terms in very loose terms about the word "context," and using that as a shorthand way to describe the situation, conditions and tools that the operators will be using during accident conditions.

I think there is a growing separation of context into two different parts. The plant context, which is what is happening in the plant, what the operators are facing, what the conditions could be, the uncertainties associated with those conditions, which is a large part of the uncertainty in HRA, coupled with what the term "task context," which is what in the past we've referred to as performance-shaping factors, performance-influencing factors and so on.

This is the tools and ammunition that the

operators have to respond to the plant context. So plant context would include is the plant in a nominal condition, is it off normal, how far is the plant going down the accident pathway. In other words, the story of what's happened so far.

The task context then is the PSFs, the training, the interface, the procedures that the operators will be using to perform their response. What I've seen in the development of the more recent methods, there's a great deal of emphasis given to the task context, but I'm seeing not so much emphasis provided on the plant context.

I think that's an area that may want further discussion, because we tend to take for granted that we almost have a deterministic knowledge of what the plant will be doing, and therefore we develop procedures based on sequences of events, the timing and so on, and yet under off-normal conditions, those sequences could be different, and the procedures may or may not be successful in capturing these alternative ways.

So I think that's an area that in the discussion of methods, and I saw in the ACRS letter, is something that needs to be perhaps discussed.

On Slide 9, a little bit about the issue of the operator inputs to the HRA methods and models. I

will say up front that as far as the IDHEAS method goes, I have not been involved in its development, nor as a reviewer. So I really don't know what the role of operators has been in the development of that method.

Other methods that have been very highly involved, the operators very highly, the ATHEANA method that you may know about, the development about ten years ago by the NRC, to capture human errors that can be induced, particularly by these unusual or off-normal plant conditions.

That relied heavily on operator input and indeed from the Seabrook plant, a willingness to use their simulator time to explore how the boundaries of the operational conditions might affect the operators. That was a critical part of the ATHEANA method. And the French method, MERMOS is built around the use of simulators and real plant operators working on those simulators as a core basis for the knowledge of what that method does.

Before I go into something that may be considered a little academic about what HRA is doing, I want to draw a distinction, and it's not in the slides. The idea came to me, as always, after you've sent the slides off on the last possible day, a distinction between HRA methods and HRA models, and I think it's an

important distinction.

The way I will do it is the models I refer to as models, that part of the HRA process that provides quantification. It's the means by which you take information about the plant context or the task context, and convert it into numbers. That is just part of the method.

And in fact if I go to Slide 11 and perhaps add some confusion by trying to draw some notional boundaries, I had previously prepared something on the world of macrocognition.

I think macrocognition can just be accepted as the way in which we understand operational processes, understanding where we are, developing the plans to respond to it, assessing the risks of alternative pathways and carrying those out. In very simple terms, that's what I refer to as the macrocognition.

Slide 11, please. So on the right-hand side of this slide, you see a box that says "HRA Models," and has inputs from plant contexts, task contexts, the PRA models and the description of the operator activities. An HRA method describes how those interrelate, so that the operators, what are they doing, how do we identify that, is it from task analysis, is it from other means, is it from the simulator?

We combine that with the knowledge of the plant context and the task context, and interact in fact two ways between the HRA models and the PRA models. The HRA quantification is just that box at the center of this, the HRA model.

So when I look at a new source of information on how HRA is being carried out, I'm trying to understand what parts that method or model or technique describes in terms of this picture, and from what I've seen, the limited information I've seen on the IDHEAS technique, it largely seems to be aimed at the modeling part.

I haven't seen, in whatever literature I've seen, understanding how the interactions with the broader PRA and the broader plant context, fit together. So that may be something we hear later. I think those were the main points I wanted to cover. I know there's a question and answer session, and my colleagues have very short times.

So I hesitate to take up the full time. So I think at that point I will finish now and pass the baton on.

CHAIRMAN MACFARLANE: Thank you. Ms. Taylor.

<u>International HRA Developments and Applications</u>

DR. TAYLOR: Thank you very much, everybody, for the invite to be here and invite to come back after the previous meeting was rescheduled. I'm working with the Halden Reactor Project in Norway, but the majority of my experience with HRA is actually from the UK nuclear industry, where I worked for approximately six years.

So that's what I'm going to focus on today with my presentation, is actually my experience of application of HRA in the nuclear industry. So on my slides, if you go to Slide 3 please. So in my experience of HRA, it's often performed as an input to the safety case, which is related to a particular plant or a particular activity.

We would perform HRA usually as part of the probabilistic risk assessment or the PRA, or else potentially also a direct input if there is a deterministic safety case, which I've often been involved in as well.

The safety case, for those who aren't familiar with the concept, it basically is a collection of documents, and it provides substantiation for the new plant or the modification to the existing plant, or the change to an activity, and it demonstrates that this new thing can be performed or can be operated within the

safety limits.

We document in the safety case how the risks can actually, or have been reduced to be a ALARP, as low as reasonably practicable, and we use a claims argument and evidence structure and defense-in-depth principles of prevention, protection and mitigation.

So the HRA fits into this by looking at the particular human error opportunities related to the new activity or the new plant, and we use the same structure then, the claims arguments in evidence, to actually provide substantiation that the operator errors are managed.

So we will usually -- and we, by we, I mean the human factors team, we're usually engaged to provide some evidence for this argument, and the HRA that we would perform would be tailored, depending on the needs of the safety case. We wouldn't perform the same process every single time, but we would actually choose how we're going to approach this at the beginning.

And the depth and the formality of the HRA that we would go through would depend on the level of risk associated with the operator actions, as defined by the PRA; the degree of novelty of the tasks and of assessment of those tasks. So if those tasks have previously been assessed in a HRA, then we would just

review the HRA and see if we need to do anything new.

Also based on the perceived complexity of the task, and that's in our opinion as HRA and as human factors experts.

Also, in terms of the opinion of the PRA people, if they think that this is a particularly complex task, then we would delve into it in more detail, and also based on the input from the plant as well. So if they think it's a particularly complex task, then we would spend more time reviewing it.

The familiarity of the HRA analyst and the plant and the tasks being assessed also play a role in the depth and the formality of the HRA. In my experience, I spent approximately five years working with the fuel storage pond operators at Sizewell B nuclear power plant.

So over time, I became very familiar with how they did things. It meant that when I was doing HRA, as time went on we would do the depth of the HRA and the formality of the process would become less, because we already had quite a body of knowledge that we were building on.

Then also we would try always to apply human factors, good practice as well, and that would influence the degree to which our HRA would actually be applied.

On Slide 5, I've tried to -- it's very difficult, but I've tried in a diagram, explain the process that we would go through in the UK, and this is fairly typical of the process that we applied at British Energy and EDF Energy.

So just very quickly, the first sort of collection of boxes at the top describes the familiarization and the preliminary assessment that we would always go through, regardless of what task we were assessing and the novelty of that task.

So we would try to define the scenario. We would review operating experience from INPO and WANO in particular, and we would go through a process of data collection, which I'll come back to in a moment, and then some task analysis and human error analysis.

Then we would, depending on whether the PRA requires or the safety case requires a human error probability, we would either quantify or we would qualitatively document our assessment. But the data collection part I've outlined in red here, and this was because of the theme of today's meeting, which is the value of operator input.

This to me was always the most important part of the HRA, with the data collection both at the site and through entities with operators and subject

matter experts. So for every HRA that I've performed, we would always, always try to go to the site, and I think about 99 percent of the time we were able to.

A site visit would include not just a plant walkdown of the area, but also observation where we could do it, review documentation on the site as well, but most importantly it was the interviews with the subject matter experts. It was really essential for us to get that operator input to our HRA, so that we could accurately reflect how things are done at the plant.

We wouldn't just assume that things are always done according to the procedures. We would want to see it as well, and it was really essential for us to actually get that input, to make sure that we are adequately reflecting the performance-shaping factors and the way things are done.

So on Slide 6, I have a statement there, which is that HRA should not be a desktop exercise, and I really strongly believe in this. I think that in order to do a good quality HRA, you have to go to the plant or to the simulator, if the plant is not possible. This is really essential.

A couple of projects that I'm involved in at the moment at Halden. I'm involved talking to a lot of HRA experts about their approach, and almost every

single one of them has said the same thing to me. You have to go to the plant. You have to talk to the operators. Otherwise, you're not really going to know what you're going to model.

So it's really important to provide that accurate information about how tasks are actually performed, information about the presence and the effects of performance-shaping factors, so to confirm or to challenge any assumptions that I may have already made.

Also we find that operators can provide input at the end of the analysis as well. So a large focus on the UK was on -- in the UK was on human error reduction, using the information that we found during the qualitative assessment, to actually try to drive improvement at the plant.

So if we've seen that a particular task, the reliability is not so good because of, for example, procedures, badly-written procedures, we can use those results to try and drive that improvement. But we need that operator input, then, to find out well, what should we do with the procedures, to actually make them better, to try and improve the reliability.

We also used the operators towards the end of the HRA, to check whether we think that the calculated

human error probability is reasonable, based on their experience, and also then for developing those recommendations for improvements.

On Slide 7, the benefits that I have found of this approach is that this detailed qualitative assessment really leads to better human error reduction. We can identify better opportunities for improvements at the plant, which was also our role as human factors engineers.

It can also assist with prioritization of recommendations. So if we found a number of areas that could be improved, it might not always be possible to make all of those improvements due to budget and time restrictions and so on. So we could then look at the HRA and see where the human error is dominated by a particular performance-shaping factor or a particular area for improvement, and we can try to use that HEP then to prioritize where we're going to focus our effort.

So that's my experience from the UK. On Slide 8, I just note that the methods that we used in the UK. While I was working there, we were using mostly HEART and THERP, which are two fairly old methods at this stage. But the UK is now also using NARA, which is the Nuclear Action Reliability Analysis, and this is a

revision and an extension of the HEART method.

They've revised the definitions of their generic task types and error-producing conditions. They've also revised the nominal values for their human error probabilities, and they've included things like an extended time factor. So to look at events that might occur over a 12 hour period and so on.

They also include human performance limiting values, and this is where if our assessment determined that actually the risk from human error was very, very low, we would apply a human performance limiting value because otherwise, it could mess up the PRA. If you've got a, for example, 10 to the minus 10 in there. It also addresses the potential for a double-counting, and also the consideration of dependency.

If you move on to Slide 9, I'll talk a little bit about our research in Norway. Basically, we were involved in a couple of projects here, including Petro-HRA, which is looking at adopting the SPAR-H method to the petroleum industry, and on Slide 10, we are also involved in some ongoing HAMMLAB simulator experiments.

Again, this is very important for us to get that operator input. So we get a lot of crews from the

U.S. and from Sweden, who come and train and work in our simulator for a week, and help us to actually run experiments on looking at performance-shaping factors, human machine interfaces and so on.

Then finally, just to wrap up on Slide 11, some of the other work that we've been involved in is the development of a HRA database, and this is something that we're working quite closely with the NRC, and also we have been involved in some of the review of the IDHEAS method, and hoping to be involved in the future testing of this method as well.

Now I've run over by almost a minute, so I shall stop. Thank you very much.

CHAIRMAN MACFARLANE: Thank you. Ms. Presley.

Industry Use of HRA and IDHEAS Development Activities

MS. PRESLEY: Thank you for inviting me. My name is Mary Presley. I'm the project manager for Human Reliability-Related Projects at the Electric Power Research Institute. I want to talk a little bit about our perspective on HRA use in industry and IDHEAS.

So if you can go to Slide 2. At EPRI, HRA research is done in two contexts. We have the HRA users group, and then we also have a broader research program that addresses method development and does kind of more

indepth research. So the goal of the HRA users group was to come to consensus on a method or set of methods, and that can be consistently applied across industry.

Towards that aim, we provide -- we have a recommended methodology, the EPRI HRA methodology. We provide application guidelines. We have a knowledge base that we maintain. We provide a software tool, which is the HRA calculator to promote consistency, we train, and then I think very importantly we provide a space for users to come together through periodic user group meetings, and share insights, share challenges and come to best modeling practices to create that culture of continuous learning in this analysis.

We also coordinate with the NRC and other key stakeholders, the owners groups, other international research organizations. Every U.S. utility is a member of our group, and we have a rising international membership. So we have that broader -- we're starting to get that broader perspective into what we do as well.

Then the broader research program looks at more strategic efforts and method developments, and this is -- it's under that broader research program that we've been involved with the NRC on IDHEAS.

So if we can go to Slide 3, so the process

of HRA, it's to identify critical operator actions, analyze them and then assign a probability, that can then be put into a system model, a probabilistic system model, a PRA, and understand how different accident sequences rank in terms of risk.

Our existing methodology we believe is -- it was developed in the late 80's and early 90's. It started developing in the late 80's and early 90's, based on a set of simulator experiments that we performed, and we believe that this methodology is fairly mature at this point, in that there's some consensus that it's a reasonable approach.

We understand where it's applicable and where it has limitations, and it's widely used with some consistency. Through focused research efforts, we've extended and augmented our existing methods for other, more challenging contexts, for fire and flood -- for fire and seismic. For example, we've also added a methodology to deal with dependency analysis.

While these sets of methods aren't necessarily as mature or not mature in the same way, as we get more experience doing these evaluations, we're bringing together the learning and refining our modeling and analysis ability.

So there are still some ongoing issues and

gaps that plague our industry. I'm not going to go into these in detail, but I have them in a backup slide if there are questions, and IDHEAS addresses some but not all of these. But I want to get to the point on the use of risk insights, and this is by and far very clear from talking to industry analysts, that this is the point of HRA, is to understand what the risk insights are.

I'm going to step back for a moment and talk about how the cycle between operators training and HRA analysts. So the methodology, while it's rooted in simulator data from the 80's and 90's, it requires, as the standard also requires, the analysts to go to the operators and get data or get the data on operations.

This is most commonly done through operator interviews. Occasionally for more challenging items, they'll be a walk-through or a talk or a simulator observation. But the analyst needs to understand the as-operated. So through that process alone, some insights come out that get fed back directly to the training and operations.

Then once the analyst goes through the HRA process and quantifies, a list of risk-significant actions and a list of time-critical actions are provided. That output from the PRA is then provided back to the Operations and Training Department for their

use.

They don't just get a list. They also get the why. The HRA tells them the why it's risk-significant, so they can then figure out what to do about it. So this is -- this use of risk insights is what's driving interest from our members to update existing models.

So if we can go to Slide 4 or -- yes, Slide 4. So EPRI got involved in this project, because we wanted to take advantage of the work that the NRC was doing, particularly to better understand the psychological underpinnings of the HRA. Operations have improved a lot in the last 20-30 years, and we wanted to have that grounding in the cognitive literature to show that in our method.

A more comprehensive understanding of potential human failure mechanisms, and we also wanted an updated approach to quantification. From what we've seen so far, we believe IDHEAS is a very positive step forward. It's addressed some of the weak points of existing methods. Particularly, it strengthened the link, we think. We'll double-check during testing, between a qualitative analysis and quantification.

It provides a more direct connection to the cognitive basis that are relevant to how plants operate

today, and then it provides clear insight on the failure mechanism and the shaping factors that inform that. I think one of the big benefits is we've taken, you know, we have the general shaping factor, but then we've parsed that into very specific questions that operators can use or that analysts can use to get that information from the operators, and better understand the context. So hopefully the risk insights then can be more actionable, clearer.

So we do think IDHEAS is a very positive step forward. We have a few cautions as we proceed, but again I'm not going to go into that. I have a backup slide if there are questions. We do understand that there's a generic methodology being developed, but we have not been part of that development process, and I think we're going to work with Shawn to see a little bit more what that's about, and how that works.

Because we are continuing to extend our existing methods to other applications. We're doing research in, you know, flooding and other areas. So it would be nice to come back and connect on the generic methodology.

In terms of Slide 5, Path Forward, we'd like to work with NRC to complete the method, finish the quantification portion and do the testing, and the

testing is very important. We need to show that this is a workable method, that it produces risk insights and the level of effort is commensurate with the risk insights it produces.

So we're going to work on the -- we are working actually with the NRC on that. But we're not waiting for the method to be complete before we start trying to use the insights that we have. We have some immediate applications of IDHEAS. In fact, we're using it right now in our dependency analysis work, to look at how failure mechanisms might propagate, and better understand dependency.

Then eventually, we'd like to put IDHEAS into our software tool and start training on it. Some real thought needs to be put into how technology transfer happens. That's one of the ongoing issues is bringing the whole of industry on board and, you know, we have new understandings, new knowledge. But getting that disseminated and constant, I guess, standard of analysis.

We need to think about how to best transition into IDHEAS, and do that technology transfer. Then finally, we have to recognize that the HRA technology will continue to evolve. It will need to continue to evolve. Operations continues to evolve.

So having a link back either to quantitative data-gathering or even just qualitative data-gatherings of experiences and having a way to reflect that in our methods and what we do, will be important.

So that's another step that we need to think about, in terms of operationalizing IDHEAS. That's all I have. Thank you.

CHAIRMAN MACFARLANE: Okay, thank you very much. Mr. Vaughn.

Experiences and Views on HRA and IDHEAS

MR. VAUGHN: Thank you, Chairman. Good morning. Jim Vaughn. I'm a plant shift manager. First, thank you for the invite today. I appreciate the opportunity to present an operator perspective on HRA.

A little bit about my background. I originally started my career coming out of college, going through the Naval training program. So there I qualified engineer of the watch, eventually a shift supervisor there.

So I was honored with the task of instructing and evaluating young sailors that became the backbone of today's nuclear Navy. It also gave me an understanding of the talent needed and where the human error first shows up in the way we operate.

Following that, I came to Nine Mile Point and licensed as a senior reactor operator in 2009, and that provided me an opportunity to apply operating experience to safely run a boiling water reactor. Also developed further insights on human performance there, as I have been deeply involved in causal analysis on human performance events at Nine Mile Point, having just completed a root cause analysis as well.

So a little about my HRA background. In order to improve the fidelity of the human response modeling at Nine Mile Point, the PRA group decided to have an on shift senior reactor operator review our HRA model.

So I was that SRO, and I gathered a bunch of insights about how HRA is applied to our risk. I also developed an appreciation for a lot of those insights, and how they could be used to improve operations, improve training and ensure we had an accurate model and prediction of human performance.

One of the opportunities I had was to support a tech spec amendment change by modeling a new operator action. I was also involved in PRA review of NFPA-805 model that's currently ongoing right now. I participated in the IDHEAS expert elicitation panel, which is one of the reasons I'm here today, and I was

also the SME for Operations in a significance determination process involving a loss of shutdown coolant at Nine Mile Point experienced in 2013.

So from my experience on HRA, I reviewed all the internal events at Nine Mile Point, and based on that review, identified several opportunities of going through there of identifying emergency operating procedure enhancements on containment venting. We identified some enhancements in our training program, based on a review of those top operator actions.

I also processed some additional procedure changes to reduce human error probability, where there were some opportunities for enhancement there.

My overall perspective on HRA, having come through all this, as well as staying within Operations, is a strong alignment within Operations and PRA group is necessary to make sure that our HRA model is accurate and we're actually using it to its true value.

The true value really is what we can glean from it to improve operations and mitigate errors. It's important that we recognize a common sense perspective of those who perform the task in the field during transients or during similar training scenarios, and as John had mentioned earlier, that context that we're talking about, the operator context, the plant

context, is something that you can't just get by looking at a procedure. So having strong tight operations really is important, to make sure that we're on the right path.

Most importantly, the exercise of steadily applying HRA methods to key operator actions should have the net effect of identifying and mitigating those barriers. At the end of the day, we have not actually been able to do anything with the methodology in terms of improving performance, and it's questionable if there's an advantage behind that.

So some of the things that HRA have to look at are the procedures, the training, design assumptions, work practices, operator proficiencies. These are all areas that we evaluate for weakness. These are also all the areas that we have to look at improvement opportunities, so that we can improve the margin we have to safe plant operation.

My experience with IDHEAS. So I participated in the expert elicitation panel along with two other Operations training instructors for various plants through the industry. Those consisted of two one week long workshops to review the IDHEAS concept.

We reviewed proposed crew failure modes, the performance influencing factors, cognitive

mechanisms and the crew response trees. We also discussed real world Operations experience for the realistic application of those crew failure modes.

So being able to talk about what the crew failure modes were, relating them back to events that we've seen in the simulator, seen in the plant, where human error occurs, was probably the most important thing that came out of those workshops.

We also provided some weighting to the performance influencing factors and estimated -- and eliminated some of the branches of the crew response trees that really would not be applicable or offer any additional insights.

So overall, I think we have a very good start on -- with the IDHEAS methodology. The comparison of performance, the performance influence factors of IDHEAS versus THERP, SPAR-H and other existing methods indicates that we do have better model of HRA on the horizon.

A key advantage of IDHEAS is that it addresses the integrated crew response, compared to a focus on the individual error drivers. So one of the things I noticed when I was going through the HRA notebook here at Nine Mile Point was a lot of it was very particular to individual failures, and didn't really

leverage how crews fail as a whole.

This is something I saw in IDHEAS method, which I think is a strong step forward in the right direction. Going forward, we need to keep a strong tie to Operations, to make sure that this really goes in the right direction we need it to, and Mary talked a lot about the testing going forward.

I couldn't agree more. A comparison of our IDHEAS results to existing HRA models to actual known performance really is the litmus test of whether or not IDHEAS will drive improvement or just provide another alternate methodology. So I'll be looking forward to seeing how that testing will be implemented.

As an example, this question was brought up by a Commissioner back in March, how would this apply to an operator action like let's say all ahead flank cavitate, right? And I want to address that straight on, right.

So I'm a not an HRA analyst, but I am familiar with it. So I went through SPAR-H, I looked at IDHEAS, and just for some ballpark numbers, from SPAR-H I looked at.

It looks like we get about 25 in 100,000 times you'll have an error associated with nominal training, versus 15 times out of 10,000 that you'll have

an error in low training. So what do those numbers mean? Are those numbers right? What does that gut feel really tell you for those of us who have seen that evolution go, and recognize the challenges associated with and the importance associated with it.

So that whole litmus test of does this really make sense. I wanted to be able to compare IDHEAS, but when I went through the draft, I wasn't able to get enough information, because not all the numbers were quantified yet to really be able to look at numbers and see if it really feels correctly.

But so in a nutshell, we're still going in that direction, and we hope that we get to a point where we can look at that, and recognize that we have an answer, which actually makes sense in the real world, and are we really able to use it and say hey, this error rate is too high.

What can we do to improve it? What training can we use, and let's use those real world examples to feed back in for an iterative process so at the end of the day, we have a tool that's worth using?

So let's see. So final thoughts is how are we going to test the hypothesis, to make sure that this method is reasonable, and the simulator data is very good. So if you look at the specific scenarios that we

run in the simulator, run through IDHEAS concept in multiple iterations and see what kind of numbers we get.

We should look at the simulated scenarios for a given accident sequence and figure out where the pinch points are. When I say "pinch points," I mean those critical moments where maybe a fast-changing parameter gets by an operator, or maybe a critical decision is made and without all the proper data analyzed an error is made.

So looking at those opportunities in the simulator, looking at the method is really, going forward, will be very important to us. Finally, one other thought I had had on this earlier in the week was having just finished up the root cause analysis back at Nine Mile Point, I was -- there's a lot of data out there that maybe isn't specific to the simulator on human error, very low level issues.

But it's there if we look for it. Perhaps there's a way to use that with an HRA going forward in the future, that you could actually analyze where error is likely across the plants on a low level, use the data to identify if our methods are working, and ultimately use that to create a refined HRA method. That's all I have.

CHAIRMAN MACFARLANE: Great, thank you.

Dr. Lyman.

UCS Perspectives on HRA

DR. LYMAN: Good morning, and once again I'd like to thank the Commission for inviting UCS to present our views, although in light of certain recent majority votes, I'm starting to wonder what the point is or if our message is getting through.

But you know, I'll keep trying. So anyway, our view on the subject of human reliability analysis in a nutshell is that we think that the subject is very important or even essential component in nuclear safety research, and the importance is clearly growing as there's increasing reliance on manual mitigating actions to comply with post-Fukushima requirements, and I think the staff briefing makes clear how far ahead they're getting on crediting quantitatively manual actions, which concerns us, that they're getting ahead of the technology.

We think that the research should best be aimed at trying to reduce operator errors and improving human-machine interface, enhancing crisis response, and the qualitative insights that these studies reveal are the most useful. But as far as developing quantitative human error probabilities and plugging them into PRAs, we have significant concerns about that.

Slide 3, please. Now if you look at NUREG-1842, which was the best practices in HRA, it says "Given the continuing importance of probabilistic risk assessments and regulatory decision-making, it is crucial that decision-makers have confidence in the PRA results, including associated human reliability analyses."

Then it says "Throughout the years, the HRA community has focused more on how to estimate human error probability, probably because this may be the most difficult, intriguing aspect of HRA." Now as a former scientist, I can see how this might be, you know, appealing.

But we're not talking about an academic exercise. We're talking about something that has real world safety implications, and they can lead to decisions that will have a real impact on people's lives. So I think you need to think hard about whether that academic inquisitiveness is really driving the subject in the right direction.

Slide 4, please. We think that aspects of PRA that cannot be well-quantified, and I say that maybe every aspect of PRA can't be, but the human error probability seems to be a major weak point, and I think that is going to damage the credibility of risk-informed

regulation as you go ahead, unless you address this, because you do not want to build on a rotten foundation, and that's what we're afraid you're going to tend towards if you don't address these fundamental issues of credibility.

Perhaps a better approach, rather than trying to quantify human error is to just admit that you can't quantify some aspects of a risk, and you're going to have the reducible uncertainties, and maybe a step function approach to human error is better than trying to come up with the continuous estimates of probabilities, the kind of step function that you've seen in the mitigated versus unmitigated scenarios in certain analyses like the spent fuel analysis.

Slide 5, please. Now from the public perspective, we don't see a lot of confidence even among the experts in this field. So I found a statement in a paper that says SPAR-H does not guarantee valid HEP estimates, which is particularly striking because that paper was written by the developers of SPAR-H.

Then we have ACRS Member Stetkar who said he believes "there's a general consensus that THERP is silly." Now those aren't the kinds of words that give a lot of confidence to the public, who may not know too much about the details.

Slide 6, and one thing I've always wondered about is the use of expert elicitation, and I think the continuing reliance or need for expert elicitation in HRA and IDHEAS is one example, no offense to Mr. Vaughn. But I think it's an admission that there's not enough data to actually come up with credible HEP estimates on the basis of statistics alone.

Now just I never really understood why if you have a subject like human error, that you think that bringing in additional human errors in the form of experts, who of course are smart people, but of course make as many mistakes and value judgments as anyone else, that that's compounding the error rather than trying to reduce it.

So I think the extending human reliability analysis to the errors made by the experts conducting the elicitation isn't the right way to go, and I think IDHEAS actually is attempting to do that. So perhaps that is a good way.

If you look at the U.S. empirical study, that really is striking in the degree of variability among different experts using the same tools, and the fact that the experts don't even understand terms of definition if you read that study.

Next slide, please. So if you just look at

some those findings, you find out that the HEP estimates and again, this was done by trying to validate a variety of models, each one used by different expert teams, against operator performance in the simulator, that the estimates themselves vary considerably from one method to another, that they vary considerably within the same method, at least in order of magnitude, difference is typical, and that the data sets themselves are being validated against huge errors, because the data sets are very sparse.

So even within three orders of magnitude between the 95th and 5th percentile, some of the guesses or some of the results of these models couldn't even find their way within that wide error. They were outside of those error bars. So that's pretty bad.

And even when the quantitative agreement is good, the study authors believe that maybe that's just a coincidence, because if they look at the underlying qualitative analysis, it didn't always -- wasn't always consistent with their quantitative estimates.

So Slide 8. So I think if you're going to apply HRA more heavily in regulatory analysis, the guidance is crucial. But if you go to NUREG-0800, you find that reviewers are only instructed that they should confirm that the modeling of human performance is

appropriate.

So here's another aspect of human subjectivity; it's the third level, is that the reviewer is going to have to review whether the experts appropriately reviewed the human errors in the models. That, I think, is taking things in the wrong direction.

So if you look at what guidance there is to try to judge if the modeling of human performance is appropriate, you find NUREG-1792, which then says that the guidance that they have is not appropriate for regulatory decision-making, and it doesn't even say it's a standard, and it's not intended to provide the defacto requirements.

So what are the reviewers -- how are they supposed to grapple with this complex subject, if they don't even have good guidance?

Next slide, please. Slide 9. It just drives the point home, is that NUREG-1842 itself says even though it's the best practices report, it doesn't provide -- it's not intended to provide any acceptance criteria for determining acceptability of PRA applications. So like I said, this is enhancing subjectivity and confusion.

Final slide, No. 10. So in conclusion, we think that it seems that large uncertainties persist in

the quantitative predictions, and even the state of the art HRAs and the empirical studies have confirmed this. I do see that IDHEAS is trying to learn lessons from these results, but again it seems to be making some of the same mistakes as its predecessors.

NRC doesn't have clear acceptance criteria for HRA adequacy, so it's hard for us to see how you're going to make the decisions to support regulatory applications. Finally, it appears that the human error probabilities are -- uncertainties can be significant to the overall PRA uncertainty, and that's another reason why we think enhanced defense-in-depth is the only way to compensate for these uncertainties.

So we weren't too happy with the Commission's decision on enhancing defense-in-depth in the context of NTTF Recommendation 1. So I will stop there and be happy to take your questions.

CHAIRMAN MACFARLANE: Great, thank you. Thank you all. Start with Commissioner Apostolakis.

COMMISSIONER APOSTOLAKIS: Thank you, Chairman. Just a general observation first. One problem that I have seen over the years with the methods is that they are too elaborate, and we have to appreciate the fact that when there is a major project being developed, HRAs -- HRA may be just a small part of it.

We saw that with the expedited transfer of fuel from the pools to the dry casts.

So the resources required to do a good job and use one of the available models like ATHEANA are not there. So people go back to simple tables like SPAR-H and so on. I'm surprised that Stetkar did not include SPAR-H in his statement on silliness.

So are we with ideas developing another huge model that nobody will use? Do you have any thoughts on that? Can we develop something simpler from the elaborate model or if you don't have an answer, that's fine. That has been the major problem so far. Mary.

MS. PRESLEY: I think that's one of the things that testing needs to show. The nice part about IDHEAS is that you build so the simulator or the -- yes, simulation experiments show the importance of a qualitative analysis. The nice part of IDHEAS is that there's a structured way that you do your qualitative analysis, and you only then evaluate the failure mechanisms if they're applicable to the task.

So you don't have to go through 14 decision trees for every single minute little task. If you decompose it correctly, the workload, we think, will be commensurate with the risk insights provided by --

that's something we want to specifically test as part of the testing.

COMMISSIONER APOSTOLAKIS: So that would be a simpler way of doing it?

MS. PRESLEY: Right.

COMMISSIONER APOSTOLAKIS: Anyway, just bear in mind the actual utilization of the model is extremely important. If you develop something that, I don't know, fits with current theories of human error but is not practical, then we're not doing much.

Dr. Wreathall, on Slide 4, you have something that caught my eye. You say -- oh, at the very last. Flooding, seismic and multi-reactor accidents. Are we showing the Slide 4? Yeah.

MR. WREATHALL: Yeah.

COMMISSIONER APOSTOLAKIS: So what is unique about, you know, the HRA for multi-reactor accidents? I mean the French are already doing it?

MR. WREATHALL: The French have a research program on the way to do it. It's not yet a method that's developed and applicable. I think there are issues of resources when it comes to multi-reactor accidents, particularly to do with staffing and sharing of resources, that may turn into risk trade-offs, that normally we think of an accident in a single unit.

You have the ability to bring all the resources, given the time available to that. But if you have distributed risks around the site, then you may have to decide am I going to put more people into one place because of something happening there than in others? So it's pushing the, if you like, the PSFs out to a further set of questions.

COMMISSIONER APOSTOLAKIS: Is it PSFs or PAFs now? Performance-shaping factors. We'll come to that in a second, then performance-influencing factors. They're the same thing, aren't they?

MR. WREATHALL: They are basically the same thing. Different people have just adopted, because they want to make a shade of difference between one and another.

COMMISSIONER APOSTOLAKIS: Why don't we go to your Slide 20?

MR. WREATHALL: Slide 20, yes.

COMMISSIONER APOSTOLAKIS: Can you show it please? You say "Not all PSFs are strong differentiators." Can you tell the Commission what, quickly what the PSF is and what this slide shows?

MR. WREATHALL: Yes. This slide and the following slide, which are meant to be taken as a pair together, come from a study that James Reason and I did

oh now 20 years ago, that looked at about 13 events for which AITs and IITs were written by the NRC.

FEMALE PARTICIPANT: What do they mean?

MR. WREATHALL: Augmented inspection team reports and integrated inspection team reports. Basically, an indepth analysis of something that was a challenge at the plant. And these documented in some considerable detail what happened at that plant.

So at that time, Reason and I looked at how plants where people did very well versus people did not do very well, judgment there. So for example, the darker shades represent the plants for which people were less successful in managing the event, and you'll see that in 100 percent of the cases, procedures were involved. The problems with procedures existed in 100 percent of the events for which problems occurred.

On the other hand, 60 percent of the cases where plants performed very well and the operators performed very well, there were problems with procedures. Procedures were not essentially a differentiating factor between good and bad performance.

So the other PSFs we looked at in this context were to do with training. Did training have issues? Were there issues to do with the organization

of the staff at the plant and the man-machine interface, HMI? And the point here was that yes, you see that the plants that had problems had generally a more frequent contribution from these particular PSFs.

On the other hand, cases where people were very successful, they still will count handling problems in their events, though it's a lower fraction. So the point partly behind this was that just simply using quality of procedures, quality of training as a way to say this will lead to good, this will lead to bad performance was not that clear. It's not that simple.

COMMISSIONER APOSTOLAKIS: But the organization, it seems to be important, right?

MR. WREATHALL: Yes. I mean in each case, each of them had a role to play. So in 90 percent of the cases where the performance was less than -- what we would judge as less than adequate, the organization of the staffing or whatever, administration was a problem. But it was also a problem in 20 percent of the cases where people did very well.

COMMISSIONER APOSTOLAKIS: Okay, okay. Thank you. I have limited time, John. Mary, from the way you spoke, I got the impression that EPRI is keeping a distance from IDHEAS. Are you participating in the development of IDHEAS, or are you just interested

observers?

MS. PRESLEY: No. We are active participants in the development of IDHEAS, and we have been -- I want to punt this back -- from the beginning of the project?

MALE PARTICIPANT: Nearly the beginning.

MS. PRESLEY: Yes.

COMMISSIONER APOSTOLAKIS: Say that again?

MS. PRESLEY: Since the beginning of the project, we have been active participants. We've been involved in the expert elicitation process, the method development. We are all in. But this comes to the -- I think maybe the reason that you have an impression that there's distance, there's the base research that we do that does research into development of methods, and then there's the user group piece, which is how the method is adopted by industry members, and the technology transfer that goes into that.

It's not -- we just want to -- we're not disavowing or distancing it from any perspective. We just want to show that just because you have a finished method doesn't mean you turn around tomorrow and it's implemented perfectly and across the board. That's the only point we wanted to make.

COMMISSIONER APOSTOLAKIS: I agree. Now it looks like your backup slides are more interesting than the main slides, both from John and you. So on Slide 9, you throw a bomb.

MS. PRESLEY: Oh boy.

COMMISSIONER APOSTOLAKIS: Barriers to applying the method. Perception that there is not consensus within NRC on acceptance of IDHEAS. Are we having a civil war or what --

MS. PRESLEY: No. This is -- maybe that is too strongly worded. Maybe the right way to describe that is we haven't heard a lot of champions within the NRC, outside of the method developers themselves, saying yes, we're going to go use IDHEAS. I think there is a lot of wait and see.

So that's in contrast too -- it shouldn't be taken by itself -- the other bullet that says basically utilities are very busy with PRA at the moment. There's a lot going on. So to get a new method adopted, there needs to be some driver, and if that driver is not because the NRC's on board and using it, then it becomes a lot harder if the NRC's not using it on their end, to fully integrate that.

So those two bullets points are kind of meant to be taken together. It's not a criticism of the

NRC.

COMMISSIONER APOSTOLAKIS: I think that's related to my earlier comment, you know. We need something simple that a user who's not an expert on HRA can use, and the users at the NRC, NRR, NRO and so on are not really experts on using an elaborate model. I mean they want something they can use immediately. Thank you very much.

CHAIRMAN MACFARLANE: Thank you. Commissioner Magwood.

COMMISSIONER MAGWOOD: Thank you Chairman, and thank all of you for coming and some of you coming again to visit with us to talk. Well first, let me sort of comment on Ed Lyman's comment earlier that he made at the beginning. I would just encourage you to always look at your participation on these panels as something that -- I'll speak for myself -- that I value, and I value your input.

I don't -- as you know, I don't often agree with you on the outcomes, but what you add to the process is always very valuable, and sometimes I do agree with you. But when I don't, I don't. But you should also know that many of the things that you and your colleagues say feed very active conversations within the agency. So it not wasted by any stretch.

So but I also have a question for you. You know, this may be actually an area where we might have more agreement than disagreement. I'm not -- I think that as we hear the conversation about HRA, there clearly is still a lot of questions and a lot of analysis and a lot of research has to be done. Your view was that it could be used to feed qualitative insights.

From what you've seen so far, can you give an example where you think the agency should be using HRA?

DR. LYMAN: Well, you know, I think it's the kinds of things that we heard from Claire, you know, where you actually -- well, let's put it this way. The parts that involve trying to use theoretical psychology to come up with some universal way that people respond, I am not too big on that.

But I think, you know, practical ways of analyzing the way people make mistakes and designing to try to reduce those mistakes, which I think there's no magic about that. But the validation aspects of these tools, I think, are crucially important, because if you don't see -- if you can't actually test your hypotheses in some close to real world fashion, then they're -- then it's hard to put any weight behind them.

I think that's a consistent theme we've

been raising in the context of all the post-Fukushima actions, that you need to have validation that is as close to real conditions as you can in an artificial environment.

COMMISSIONER MAGWOOD: Well you know, in a way your comment just sort of raises something that I observed as I was listening to the panelists. Each of you spoke of the application of HRA in somewhat different terms, you know. I think I heard Mr. Vaughn talk about improving Operations. I mean that's how you view its use in your company.

I think Dr. Taylor mentioned improving -- basically putting it in the context of improving human effects, improving procedures, improving performance, and Mary, Dr. Apostolakis just calls you Mary, so I call you Mary; I hang around him too much. You're the one person that I liked integrating HRA analysis into PRA models.

I wondered -- I just wanted to ask the panel, this side of the panel, because I think Dr. Lyman's views are clear. Is everyone in agreement that we should be integrating HRA into larger PRA models, or should we look at HRA as a stand-alone tool unto itself for specific applications? Sort of start with Mr. Vaughn and work our way down.

MR. VAUGHN: I think there could be advantages to integrating the PRA model. The major advantage that I spoke to, though, is the exercise of going through HRA and identifying weaknesses in operator actions, things that are especially important to us, that we're successful in gathering.

Those insights are -- should be the first priority. Integration of the PRA model, improve accuracy downstream as a whole could be a secondary advantage.

COMMISSIONER MAGWOOD: Okay. Mary, do you want to comment?

MS. PRESLEY: My inclination is to say yes, because when you're making -- in terms of integrated decision-making, when you're making a decision, you're balancing different aspects, and the mechanical systems are one part, and yes, maybe we have better data on it than we do for human performance.

But human performance is such a big part of how a plant runs. I don't think that you can separate the two. I think it would be artificial to separate the two and create more, I guess, maybe false -- it will create a different impression that's not true.

I do recognize that probability -- I mean I've heard HRA called the dark science or the black

magic, right, and it's true. There's some squishiness to the quantification part, because we don't have a lot of hard data for these things.

But to be able to focus on the relative rankings and the insights that they provide, we need a tool that we can look at these things systematically, and it's the only tool that we have.

So the next question is if we don't use this, what do we use, and we have other aspects. Like Claire mentioned the defense-in-depth and programs to make sure that organizations are, you know, have a good safety culture and other programs in place to shore up the residual risk where we can. But we still need a tool by which we can make decisions, and this is kind of what we have. So I don't think the focus on numbers should be a killer of PRA, or HRA and PRA.

COMMISSIONER MAGWOOD: I appreciate that. Dr. Taylor.

DR. TAYLOR: Yeah. I think, you know, from my perspective, as I said in the presentation, the real strength of HRA is that it gives you a stick to wield, to show how much human error can actually dominate within a PRA.

So I think it really is important to integrate the two and, you know, my experience prior to

that is that if you're going in and trying to assess situations and assessing them as a human factors expert, it's very difficult to get the attention from the right people, to say -- to demonstrate how important this is.

By putting it in the PRA, you can show, using numbers, how much of an effect it has. The numbers aren't perfect and, you know, the methods that we use aren't perfect. They are human error probabilities, they are estimates. But if you have a good analysis behind them, you can have a good degree of confidence in how much this error may dominate, and therefore you can use that to drive the improvement.

So from that point of view, I think it is incredibly important. Of course you also have the flip side then, where you may see an issue that you think is quite important, but actually it doesn't dominate the PRA sequence. So therefore, how do you actually get the resource and the budget and so on to drive those improvements.

But I think that's the potential downside of it. But I haven't seen that too often. I see that usually it's quite a good way of actually, you know, shining a spotlight on the human side of operations.

COMMISSIONER MAGWOOD: Thank you.

MR. WREATHALL: Yes. I think in part the

question comes to both HRA and PRA, and that is the reason why it's being done. I somewhat simplistically break out three different reasons why you might do PRA and its human component. You simply want a probability number. There is a quantification need; a number is needed.

The second and perhaps more useful thing is that from a human point of view, you're trying to improve or optimize the design of the human interface, or the procedures or the training. So it's not just a number. It's a I need some knowledge about the situation and what I'm gaining from it. Then the third, which is the bigger question, and that is what is the integrated plant safety that takes account of the potential for human errors, and that really does involve a complete combination of HRA, PRA, whereas the optimization part for humans could be a narrower thing. I think it connects very closely to Dr. Apostolakis' first question about is there a simple method, is there a much simpler way of doing this. It seems to me that you can develop relatively simple methods that address different issues.

But if we're looking for a single big HRA box that will do all of these for many different conditions in plants, we are going to finish up with

complicated models. The first step in the ATHEANA method is what is the purpose of this analysis, and can I select just a narrow set of tools and methods that address that, that reason.

Whereas when we talk about a comprehensive set of methods, they're really never to become complex, because they're trying to answer many different questions, not all of which are relevant to this particular issue. So that's my response.

DR. LYMAN: Let me just clarify something. I think our main concern is the development of absolute values, you know. You calculate a core damage frequency or you compare it to the safety goals, and if you're basing that on an absolute value without quantifying certainty, that's the problem.

But what I'm hearing more is sure, if you use that to study the relative importance of various factors, then those uncertainties are, you know, cancel out to some extent. So again it's the -- so I don't think we have a problem with using it to study, you know, the relative changes in risk as opposed to just putting, plugging in these absolute values.

COMMISSIONER MAGWOOD: That's a good comment. I think you'd find a lot of people agree with that. All right, thank you. Thank you, Chairman.

CHAIRMAN MACFARLANE: Thank you. Commissioner Ostendorff.

COMMISSIONER OSTENDORFF: Thank you, Chairman. Thank you all for your presentations. I'm going to make a couple of quick comments before I get into questions. John, I appreciated very much your kind of capturing the worldwide perspective on methods being used. That was very helpful.

Claire and Mary, I appreciate your coming back. I, like Commissioner Magwood, have benefitted from meeting with you a few months back and I appreciated your comments, in particular the identification of complex tasks. The comments you made was important from my experience. Mary's use of risk insights, that terminology, and both of your reliance upon interviews with operators, I think, was right on the mark.

Jim, I appreciate your operator presence here. It's really important. I know it's been echoed by the people to your right and to your left. I think your comments on the containment venting strategy is a potential area to explore. I'll come back to that later on, and your shift supervisory experience at Nine Mile Point is very crucial.

Also as a former Navy guy, I appreciate the ahead flank cavitate. In the 1990's, I think, I had a

chance to shoot a 480, Mark 48 Adcap torpedoes. You know, as a commanding officer of a submarine or in charge of commanding officer training for Atlantic Fleet.

But a key part of that was torpedo evasion, and so the head flank cavitate example you used was a great example of what, I want to use your term, integrated crew response, as to how to conduct an operational event in less than one minute, that involved coordination between the reactor operator, throttleman and the engineering officer of the watch, in a very dynamic environment. I've seen it hundreds of times. I thought that was a great example.

Ed, I appreciated your comments. I want to first agree with Commissioner Magwood's commentary on how we value your participation, and but I also appreciate the fact that you made a statement expressing your concerns. I think that's important for us to hear that, and I was not surprised by your comment, but along with the rest of the Commission, I know we all value the UCS role, and perhaps you do, as Commissioner Magwood noted, have a greater influence than perhaps you think you do.

I'm going to start out with the comment you made on the HRA topic, and that was I agree with you on the qualitative use of the HRA principles. I'm not

opposed to quantitative. Mary and Claire and I discussed this in my office at some length a few months ago.

But I think that certainly I think your statement was that perhaps the HRA studies are most useful in providing qualitative insights. I agree with that, and I'm going to provide a contextual example to frame a question for all of you in that area.

So I go back to 1985, when I was an engineer on John Marshall, a slow attack submarine out of Norfolk. The Naval reactors program sent every submarine its own, I think it was called a primary plant response demonstrator. It was a box about that big (gesturing), that long, that high, and it was the first simulator that I ever saw used in the Naval reactors program.

On submarines, you did all these actual drills. You did SCRAMs, flooding, stream line rupture casualties. All those things are actually done on the plant, as opposed to simulators. But because of concerns on the operator ability, and primarily the reactor operator ability to recognize a primary coolant leak, and to discern the parameters, is this a slow leak, which is X inches per minute, that still is classified.

But X inches per minute pressurized level

drop from a fast leak, which has a greater number, and there's different sets of actions from both those kinds of leaks. You're nodding your head. You know what I'm talking about.

So primarily to help provide better operator awareness and to train the operators in detection and recognition, Admiral McKee, when he was head of Naval Reactors back in the early 80's, simply you know, mandated we use these, and they were very effective training tools.

It was really trying to look at the HRA aspects of how hard it is to determine, when you're watching this gauge, among 20 gauges in the maneuvering room on the reactor plant control panel, this level indication coming down to a certain rate would determine what operator action you should be in, fast or slow leak.

Another example, again I'm setting it up for question here, was you know, as a result of the loss of the USS Thresher back in the 1960's, the Naval Reactors Program developed what's called a fast recovery startup. The details of that procedure are classified, but basically it was an emergency startup.

And as part of that emergency startup to restore reactor power, to restore steam to the turbines to be able to drive the submarine to the surface in the

event of a flooding casualty, you had reactor start being conducted in a very short period of time, with very high startup rates, with high heat-up rates.

So the integrated crew response piece that Jim's mentioning required great coordination between the Reactor Operator, the Throttleman and the Engineering Officer of the Watch. So that was another example I thought was relevant to the use of identifying potential areas, with Dr. Lyman's comment on qualitative factors requiring a lot of training and reinforced training.

So those Navy nuc examples, I wanted to see if there are any operator plant examples from a training or procedural standpoint, that you've identified as needing work or areas of potential application. I know that Jim mentioned containment venting. I believe that Ed may have a -- I'm going to ask him a question about manual operator actions in Fukushima.

But I'm trying to understand what have you seen so far from your experience that indicate areas for improvement apply your HRA experience, to help focus on procedures or training? I'll start from the left and we'll go down the line there.

MR. WREATHALL: Thank you, yes. The concern I have in trying to answer the question is I'm

going to try avoid answering the question, and still trying to give you some useful answer.

This issue that I keep raising about plant context is very important, because it represents the potential divergence between what the designer assumes will happen in the plant at any given event, and what really is going to happen, and the typical examples of plant context that affect that are failures in other ancillary equipment or something else going on in the plant, whereas the designer, when he's writing the procedures or developing the maneuvering room designs and so on, is assuming that this is the only thing that people focus on.

So what I have seen in plants and what we found when we did the simulation trials with ATHEANA is how much does the plant have to be away from that nominal designer's mind assumption about what's going on, before the repetitive training in fact is going to capture people into something where they really should be questioning it.

I haven't seen that much in the way of application of that concept into training. I mention in my bio that I'm working in a field called resilience engineering, which is sort of a parallel but somewhat different from PRA. Its purpose is to --

COMMISSIONER OSTENDORFF: I'm sorry. I'm going to run out of time here. So I got your point. Thank you, and we'll go down the line here.

MR. WREATHALL: Yeah.

COMMISSIONER OSTENDORFF: Sorry, thanks. Claire.

DR. TAYLOR: I think in a nutshell the area that still concerns me is more the issue of dependency between events or between potential human errors, and how we model that in HRA. That, I think, is one of the areas, and we've discussed this before. I think that's still one of the really big gray areas. So how one event influences the next and the next, and causes the error. I think that's the part that HRA needs to be focusing on more.

COMMISSIONER OSTENDORFF: Okay, thank you. Mary.

MS. PRESLEY: I think we've seen a lot of improvements in fire. I think that's one of the big success stories. I think when we get into some of the other severe external events, we're going to have to start looking at, I guess, decision-making and command and control. Main control room abandonment is one of the areas where command and control comes up.

But it comes up in all sorts of areas. But

that's one area. It's in my ongoing issues slide.

COMMISSIONER OSTENDORFF: Okay, thank you. Jim.

MR. VAUGHN: A couple of things, that going through our HRA notebook, we identified areas where, for example, we have these things called recovery steps. If you have a step following an action that says "verify this parameter is good," you get credit for that, to basically say hey, you have another opportunity here to catch something you previously missed.

So going through the HRA notebook, I identified various procedures where, you know, just adding that step in here was something we could add on, to help mitigate risk in an accident. Is the operator just supposed to validate that anyway? By actually putting a procedure when, you know, the stress levels are high, is really one good way that we can use to improve it.

And that's using, you know, CBTM, previous HRA methods. But the idea is right now it's still early on. I don't know what the full scope of that would be in the end. But the idea that we could look at how crews could fail and how crews or pinch points associated with the crews and put in, you know, the equivalent to recovery steps there in the training process and our

procedures and use that to improve, I think could offer a lot of advantage going forward, depending on how we implement this.

COMMISSIONER OSTENDORFF: Thank you. Ed.

DR. LYMAN: Well, I just -- just focus again on, you know, these critical paths that you're building into, you know, post-Fukushima response, and the realism of some of them.

For instance, the flooding has come up and I'm very interested in seeing how -- how those flex strategies are going to be developed in a way that is really credible enough that you can have confidence in approving them, like having to move equipment in advance of a rapid -- a rapidly advancing flood in enough time. So you know, that's one separate aspect which I think needs to be considered.

COMMISSIONER OSTENDORFF: Okay, thank you. Thank you, Chairman.

CHAIRMAN MACFARLANE: Thank you. Well thank you all for your presentations. I'm struggling with how meaningful any of this is. So you know, I'm struck by some of your statements. Ms. Presley said if we don't do this, then what do we do to analyze? I've heard that before, and Mr. Wreathall said we need a number. Do we?

You know, if your number isn't meaningful, then what value is it? And I'm worried that maybe this distracts from actually more truly meaningful ways of ensuring safety. So I think we really need to be very mindful of the limitations of the methodologies that we rely on.

I'm struck by two omissions from the discussion that the four of you had for the most part. Ed talked about this not explicitly but implicitly. The first is a discussion of uncertainty. None of you mentioned uncertainty. It seems to me that the uncertainties are enormous here, and I'm interested in how you quantify them.

You know, Mr. Vaughn talked about something feels correct. I think that's fascinating language. I think the language that people use to describe -- you know, all the language you've been using here is fascinating and worthy of a good social science study, which one day I will conduct, but not today.

So I'm curious, very briefly, if you would discuss just how you quantify uncertainty. Let me just go down the line real quickly. Quickly, because then I have another question. Actually, I have a whole lot of questions.

MR. WREATHALL: Okay. I just want to be

clear that I wasn't saying we do need quantification. I said it's one of the three reasons why people do perform PRA and HRA. It may not be the most important one, but people do use it for that.

CHAIRMAN MACFARLANE: Right, yes.

MR. WREATHALL: And therefore related to quantification, and its need or not, the uncertainty to my mind, and I keep coming back to this same point.

It's in many ways the uncertainty about the inputs that go into understanding the situation we're going to analyze, work that is beyond the scope of this discussion, is an area that I'm involved in, that is actually trying to represent, as best we can, the uncertainties in just defining what the situations will be that operators face, and how that would play out in not just numerical uncertainties, but in uncertainties in the pathways they may take.

So I don't have a good answer to the immediate quantification of uncertainty, but I don't think that's the driving issue right now.

CHAIRMAN MACFARLANE: If you have a model, the model is useless unless you understand the uncertainty associated with the result. It is useless, and if you have not quantified that uncertainty, throw it away. You've wasted your time. Go ahead.

DR. TAYLOR: The way that we've dealt with uncertainty in the UK is it's incredibly difficult to quantity. So the best that we could do is to document our uncertainty very clearly, and try to quantify anyway, and then review that uncertainty as time goes by.

So if it's on a larger project, some of the ones I was involved in were over five years, they would constantly go back and review what we had documented, to see do we know anything new now that changes that.

If not, when it comes to the end of our analysis, the end of our safety case, it's documented. So when that safety case gets reviewed again, at least it should be clear to the next people coming in looking at it what we based our analysis on.

CHAIRMAN MACFARLANE: The value of the safety case.

DR. TAYLOR: Yeah. So that was the best that we could do.

MS. PRESLEY: Very similar to what Claire does, we document the source of uncertainty. We do put an error factor on these numbers. There's a rule that we use. But I think most importantly, I just blanked out. Sorry. Give me a second.

CHAIRMAN MACFARLANE: That's okay. Mr.

Vaughn.

MR. VAUGHN: I shared a similar concern when I was sitting the IDHEAS panel there, of saying well, what's the uncertainty? How accurate is this number? It was difficult for me at times, even on the panel, was you know, thinking of anecdotes, thinking of examples of how this fits in, and in one case where, all right training is of the utmost importance.

No problem; we'll always address this; but at other times, well maybe not and how do you really quantify that, if you ask me is it 1 in 10,000 or 1 in 50,000 or 1 in 100,000.

Humans don't have that gut feel, so to speak, to be able to really know if that really makes sense. We have a very limited scope, especially when we're talking about accident space. Now if you go look at more every day kind of minimal errors, and expand an HRMF to include every day minimal errors, I think you have a much broader set you could actually pull from, and get real uncertainty.

But when you're talking accident sequences that never happened, even though they happened in the simulator, it's not the real plant. It's not the same. The operators are under a different kind of pressure, and it's only a resemblance of what we're actually

trying to model.

CHAIRMAN MACFARLANE: Excellent segue to my next question, which is on where you get your input data. Now I understand from the discussion so far, a lot of the input data or the vast majority of it comes from simulators. Again, it's a simulation. It's not reality, which is your point.

So why aren't you -- why aren't we talking about reality? Okay, there are real accident experiences, okay. We have TMI. We have Fukushima. You could compare Daiichi to Daini responses. There are other less significant accidents that you could look at in the nuclear realm, and you can go beyond that.

I think there is a set of unfortunate data out there that -- in the sense that it was bad news for the people who experienced it, where you know, this good data doesn't support the value of training. For instance, the recent ferry accident in South Korea. The Italian cruise liner accident last year, where you had trained crews who basically fled, or the captain anyway fled.

You have, you know, the Air France flight from Brazil, where the pilots didn't believe their instrumentation. You have, you know, the behavior of soldiers in World War II, where a significant percentage

of them didn't -- actually did not use their weapons they were trained to use them.

You know, there actually is a lot of actual data out there that you could use and input. Do any of you work with social scientists, sociologists, etcetera, to collect data?

MR. WREATHALL: I have been working particularly with James Reason in the UK, who has developed handbooks of those kinds of data, both from the most trivial level of error up to performance data, railway systems, health care, and there are two HRA methods. To a large degree NARA is based on the digestion of those kinds of data, not just from the nuclear, but from other fields, and there's a German method called CAHR, that also is built on experience data in the German plants.

So there are actually methods that are out there that are using precisely that approach. The problem is one, how does that -- those data connect to the severe accident situations that the PRA is trying to model, and we're back to the uncertainty issue then.

CHAIRMAN MACFARLANE: Right, and what Dr. Taylor mentioned, which was these issues of dependency. There's a social scientist named Charles Perrow who described normal accidents, where you have these can't

imagine or unexpected situations, where you have tightly coupled systems that produce these accidents, and TMI, Three Mile Island was one of them that he used as an example.

So how do you -- how can you actually test these models? Have you tested any of them? I'm talking about the validation and verification piece.

MR. WREATHALL: To a limited degree, and I think we have to admit it is to a limited degree, and again, in the ATHEANA method, we came up with a working model. We took it to the Seabrook simulator and worked with the trainers to see if indeed what we hypothesized would happen.

CHAIRMAN MACFARLANE: But it's a simulator.

MR. WREATHALL: Again, taking a plant to core melt --

CHAIRMAN MACFARLANE: No, I wouldn't do that. But I think you can -- as a general rule, I wouldn't do that. But I think that you could try to apply the models to, you know, proto-accidents if you want to call them, that are situations that develop in plants, you know, which happened on occasion.

MS. PRESLEY: I mean we have looked at retrospective analyses, if that's -- I mean that's one

way. It doesn't -- it can't test the quantification part, because we don't have a denominator and a numerator. So from that, I mean we do look at facts and data out there. Those do inform our methods, but only in a qualitative sense.

And then I was wondering if I could just take a moment to address your question on uncertainty. I think understanding how the HRA happens might help alleviate some of the concerns with the uncertainty. As Dr. Taylor mentioned, you quantify at different levels, depending on -- you put more effort into it if it's more important.

So a lot of the analysis starts with put in a 1.0, and if the model tells you it's significant, then you start looking at it in more detail, and you do more work based on its risk significance, to understand the story and the detail and the context.

CHAIRMAN MACFARLANE: Yes, but if the model is incorrect to begin with, you're following, you know, an incorrect trail?

MS. PRESLEY: Well, the model is right.

CHAIRMAN MACFARLANE: So you assume the model is correct?

MS. PRESLEY: There is model uncertainty, and we do look at that in PRA space. There's guidance

on how to look at uncertainty, model uncertainty and parameter uncertainty.

CHAIRMAN MACFARLANE: Where does that guidance come from, the Gods?

MS. PRESLEY: EPRI-26511 and NUREG-1855. But it's how to look at key sensitivity studies. Again, you have to go to the context of how you're using the PRA. So you identify your key sources of uncertainty and then you do sensitivity studies to understand how that would influence your decision.

You don't just do a PRA just to do a PRA and come up with a magical number. That's not -- and I think you appreciate that. But you really have to talk about the specifics of the decision that you're making, and understanding uncertainty in that context.

CHAIRMAN MACFARLANE: Yes, I know.

MS. PRESLEY: So the fact that we can't put data, you know, put large uncertainty bounds on the data and put it in our model, maybe that's not the most useful approach. So we break down the question of uncertainty in different pieces, and then look at the pieces as we can. That was in part was how Claire described it.

CHAIRMAN MACFARLANE: Okay. I'm way over my time. Thank you. Commissioner Svinicki.

COMMISSIONER SVINICKI: Well thank you all

for your presentation. I'm not a practitioner of HRA, so both from this discussion and in preparation for this meeting I learned quite a bit, and I do agree with a number of my colleagues and, as a matter of fact, all of you as experts, that there are a lot of challenges here.

But I don't see that as a reason, you know, to give up. I think this is a very worthwhile area to continue to try to advance the state of our knowledge. I am maybe a little hung up on some of the same areas that my colleagues are.

I do want to note, Chairman Macfarlane didn't make reference to this, but maybe it was the source of developing some of her questions, is the Advisory Committee on Reactor Safeguards, in their review of human reliability analysis models, spent a quite a bit of their letter report on an integrated assessment of uncertainty.

So it was something that the ACRS pointed out as well. They said the topic of uncertainty is afforded only cursory attention in the IDHEAS draft report, and they go on to argue for greater reliance on expert elicitation processes. I think that's one of the strategies that they recommend to the NRC staff to make heavier use of.

I also acknowledge some of the difficulties when we look at modeling human behavior. That seems like one of the very big challenges. But I agree, as others have noted, is that at least I think there are analogous sets of data.

It occurs to me, you know, that the U.S. military puts groups of trainees through standardized training and exercises, and there is, I think, some data monitoring of performance of troops in the field.

So I don't know if that's a source of any types of data. But it did appear to me, again as a non-practitioner of HRA, that there's a lot of discussion of human error. But in agreeing and aligning myself, which I do with the point that any model is going to have to be tested against real world experience, it occurs to me that that needs to cut in both directions.

So I became in my mind kind of hung up on this question, which is if one -- and it's a non-nuclear example, which was I think actually helpful sometimes to use something that's not a severe nuclear accident. But if an HRA practitioner used any of these models to look at an airplane crashing into a high rise building in New York City, and was trying to make assumptions about the behavior of New York City firefighters and

first responders, would the result be that there would be less -- more civilian deaths and less firefighter loss of life in the buildings as they were collapsing?

Because would the assumption be purely logical, that with their knowledge of structures and fires, they would realize when the building was about to collapse, and there would be no room in the models for a demonstration of human behavior that is extraordinary or heroic? Is there no way?

So it seems to me, you know, if a model is going to be compared to real world experience, real world experience tells us that in addition to some percentage of human errors, there are going to be some fraction of human beings whose conduct or behavior would be extraordinary and outside the norm.

It's not all human beings, but some fraction, because we routinely find that in emergency situations. So do any of these models, can they accommodate at all the fact that in real world situations, there would be some extraordinary conduct.

Frankly, I don't know how you would model it, but I ask the question simply because I'm not familiar with what's embedded in these models. Would those New York City firefighters just be standing on the sidewalk and watching the building collapse? Is that

what you assume?

MR. WREATHALL: There has been quite a bit of work done that I don't think is formally incorporated in HRA models, but humans as hero, and the fact that people will take on the role of going way beyond what you would expect a rational, normal person is, and you uncover that by following them, understanding their culture, and seeing how they've behaved in very similar situations that perhaps weren't as catastrophic.

In my slides, I refer to the work of Gary Klein, who is a psychologist who has done a tremendous amount of work in understanding in military settings, in firefighting settings, in rescue settings, how the hero comes about. Now we haven't taken advantage of that. It's certainly in the nuclear power plant PRA formal settings, because we focus on the bad side, if you like.

I think as we look to Level 2 and Level 3 type PRAs, where it's an area that heroic action may play a role, we might want to consider how to add that. But there's nothing in the modeling right now. But there is the qualitative understanding of how people can become heroes and take on those roles.

So it isn't something we've neglected. It's something that in the scope of PRA and HRA in

nuclear plants we've had no need to push that far yet.

COMMISSIONER SVINICKI: Okay, thank you. That's helpful. I don't have any other questions.

COMMISSIONER APOSTOLAKIS: Just a comment which is somewhat related to what Commissioner Svinicki just said. There is another source of uncertainty, and we do have data on those, where the operators came up with very clever ways of handling an accident that as not in the procedures.

This is documented fact. I think it goes back to the Brown's Ferry fire, as I remember, where they used the firewater to cool the reactor. But nobody was telling them to do that, and that is completely ignored by these models. The fact that the operators may do something smart is not there. So that's another source of uncertainty which is a good uncertainty, okay.

So and I get the sense that, you know, all these discussions of validation or whatever, the conclusion should not be to throw these models away. And again, even with quantification, you start thinking okay, I'm not going to quantify. The probability is 1 that they will make mistakes.

Well, we can't live without that. It's not 1. We know it's not 1. So the big question is how far down do you go, okay, and I'll leave it at that. Thank

you.

CHAIRMAN MACFARLANE: Okay. We will -- thank you again, panel. We will now take a five minute break, and then we'll have the NRC panel.

(Whereupon, the proceedings in the foregoing matter went off the record at 10:50 a.m. and went back on the record at 10:56 a.m.)

CHAIRMAN MACFARLANE: Okay. Ready? All right. Now we will have the NRC panel. I'm going to turn it over to Mike Weber, our Acting Executive Director for Operations.

MR. WEBER: Good morning, Chairman and Commissioners. It's a pleasure for the staff to appear before you today. I would just add before we actually get into our presentation, we very much appreciated the presentation of the last panel. I think you had a healthy, diverse set of views, but they were all very well informed and I think that contributes to the work before the agency.

We rely on people to accomplish safety and security when it comes to the safe and secure use of nuclear materials and facilities. So our analysis, our understanding of the contributions that their performance makes to safety and security is very important to us.

I think the information before you makes a compelling case. We've made a lot of progress over the years in this area, and so we're proud of the progress that we have made. But we certainly recognize there are many challenges that remain before us, and we've got a very dedicated staff focused on, how do we make progress on those challenges, and how can we continue to use human reliability analysis as a tool in our arsenal to contribute to safety and security.

For our group today, we're going to have Rich Correia. Rich is going to talk about the role of human reliability analysis and our regulatory framework. We have Dr. Sunil Weerakkody, who is going to talk about, how do we actually use human reliability analysis in regulating nuclear power plant safety. And then Sean Peters is going to follow up with a more detailed review of the method that we have developed over the years, the scientific basis for that method, and the steps forward as we proceed.

So with that, Rich?

MR. CORREIA: Thank you, Mike. Good morning, Chairman, Commissioners.

Let's go to Slide 3, Introduction to HRA. Yes, thank you.

As you have heard before from the other

panelists, human reliability analysis addresses the questions, what actions do humans need to take and how likely will they succeed or fail at performing those actions?

The information from those questions become an integral part of the probabilistic risk analysis that is used to evaluate the consequences of human errors and a contribution to public risk. Human reliability analysis is important, as you've heard. Human errors can be significant contributors to events and actions, not only in the nuclear industry, in many industries.

As part of our regulatory decision processes, human reliability analysis can provide a description of the human contributions to risk to the public and, thus, can be used to identify ways to reduce risk through orders, rules, guidance, and information.

Without human reliability analysis, probabilistic risk analysis would lack insights into the very large influences that human reliability has on overall risk, which could result in focusing resources on less risk-significant areas. Probabilistic risk analysis treatment of human reliability needs to be similar enough equipment reliability that the probabilistic risk analyses can produce balanced risk

insights into what aspects of the facility are risk-important.

Next slide, please.

Human reliability analysis is an important part of our regulatory decisionmaking processes, such as the bases for orders, rulemaking, oversight, licensing, generic issues, events analysis, and research products, such as the Level 3 PRA.

For example, a complicated event at the Robinson Nuclear Plant in 2010 that involve equipment malfunctions, two fires, and failures of operators to diagnose plant conditions, and probably control the plant, contributed significantly to plant risk. The operators took actions to bring the plant to a safe and stable condition, and the event did not adversely affect the health and safety of the public.

Our human reliability analysis of the event found that weaknesses in operator training, emergency operating procedures, and command and control in the control room were important contributors to the overall change in plant risk for that event. For that event, we gave the licensee seven findings ranging from low to moderate safety significance to very low safety significance.

The Robinson licensee took extensive

corrective actions to improve operator performance to prevent similar events. These corrective actions were made using human factors engineering principles to improve procedures, training, control room management.

Dr. Sunil Weerakkody's presentation will include how the staff uses HRA to address the risk significance of this event as part of the reactor oversight process.

We also used information -- we also issued an information notice about this event to alert other licensees of the problems Robinson faced, so that they could evaluate their own programs to avoid similar events. Other examples of where we use human reliability analysis was the consequence study of a beyond design basis earthquake affecting a spent fuel pool and the ongoing containment filtration strategies and regulatory analysis.

Next slide, please.

The main focus of our briefing, as Mike said, is on the results of the staff's efforts to develop human reliability analysis methods. We recognize that HRA is a very challenging -- is very challenging. And as a learning organization seeking to continually improve our methods, we have made significant progress.

The integrated decision tree human events

analysis system, or IDHEAS, is the HRA method that the staff developed for analysis of reactor internal events at power. During the development, the staff had positive interactions and feedback from ACRS, extensive collaboration with the staff and external stakeholders, many of which were at the panel here previously, and I'd like to take this opportunity to thank them for their voluntary efforts to help us develop the IDHEA methods. And they will likely continue to do so.

This improved method uses best features from other existing methods, has enhanced capabilities, and was built on state-of-the-art technical basis. The generic method is also under development and can be tailored for various applications, not just reactors at power. You will hear more details about these methods in Sean Peters' presentation.

Now I'll turn to Sunil Weerakkody for his presentation on the regulatory uses of HRA.

DR. WEERAKKODY: Thank you, Rich.

Next slide, please.

My name is Sunil Weerakkody. I'm the Chief of the PRA Operations Support and Human Factors Branch. I want to use the next 15 minutes to discuss the importance of human reliability analysis in our decisionmaking. I also want to make some remarks with

respect to how some of the research that the Office of Research is conducting is very relevant and will be useful to us.

There are a number of areas in reactor regulation where we use human reliability analysis to make significant impacts on decisions, and I'm going to mention three examples. We use human reliability analysis to determine the significance of inspection findings as part of our reactor oversight process.

We use human reliability analysis to the risk-informed license amendment request. We may use human reliability analysis in the rulemaking process as part of the reg analysis. In addition to these applications, I want to point out a few areas where a licensee may use human reliability analysis and its insights to enhance plant operations.

Next slide, please.

I'm going to use the event actually to mention -- the event at H.B. Robinson Unit 2 to further elaborate how we use human reliability analysis in the risk-informed reactor oversight process. As Rich mentioned, the event at Robinson involved equipment failures, fires, failure of operators to diagnose problems at the plant. One performance deficiency that we have to analyze as part of our oversight process is

the operator failing to follow procedures and the deficiencies in the command and control functions in the control room.

We used an HRA method called SPAR-H, which you heard frequently, to estimate the risk significance of this deficiency. We selected contributing factors -- we call them performance-shaping factors -- to evaluate the appropriate increase in the failure probabilities using the published guidance containing the SPAR-H methodology.

We did sensitivity analysis as necessary. Then, we applied expert judgment, as appropriate, to increase some failure probabilities to reflect the performance deficiency.

Let me elaborate a bit on that. During this process, using the guidance in SPAR-H methodology, we changed probabilities of some failure of some operator actions by as high as an order of magnitude from the nominal value. We selected these values using expert judgment as appropriate. We determined that the risk significance of this performance deficiency is wide, though we call it low to moderate. Had we made only minor adjustments to these failure probabilities, the finding could have been green.

One of the things I want to do here -- it's

not necessarily in my prepared remarks -- is in the previous speakers there was a lot of discussions with respect to some of the uncertainties with respect to the human error probability. I want to make sure that when we use SPAR-H, the numbers we calculate is the starting point for discussions. In other words, in this particular exercise, we don't just plug in the numbers and run with it and make the regulatory decision.

When we do that initial calculation, it tells us exactly what are the key areas that could influence the answer. And, if necessary, we would -- I would send some of my staff to talk to the operators, talk to the licensees as necessary. So I think the advantage of SPAR, in spite of some of the weaknesses that you pointed out, which means it does not give a guaranteed number, is it clearly helps me make that high-quality regulatory decision by focusing my staff to dig into the right areas. I just wanted to make that point here.

And I just gave you one example of how we use human reliability analysis in reactor oversight. Human reliability analysis, as you already know and reiterated, is not an exact science. However, as demonstrated here, it provides a very powerful tool to us to make meaningful distinctions when we make

important regulatory decisions such as the ones that I just talked about.

In fact, more often than not, human reliability analysis becomes one of the critical inputs to the decisions in the reactor oversight process. That is because operator actions in some form are a part of the response in many event sequences.

Next slide, please.

Now I'm going to take an example of an operator action whose reliability may make a significant change in the regulatory decisions pertaining to risk-informed licensing action. For this discussion, I am picking a very timely topic. I'm selecting the reliability assigned to control room evacuation in fire PRAs.

As you all know, a number of licensees have done fire PRAs, and some of them are already performing fire PRAs. When they perform fire PRAs, one of the things they need to look at is the sequence where the operators may have to leave the control room or evacuate the control room.

They may have to do it for two reasons. One, there may be a fire starting in the control room, and the shutdown systems may have -- even though unlikely, may not have worked. Or there could be a

scenario where a fire is in a different area of the plant impacting the operator's ability to control the plant from the control room. In either case, the operators must leave -- evacuate the control room, but every plant has remote shutdown panels from which they can control the plant.

Now, the human error probability that we assign to this particular probability can be very critical in our decisionmaking. In fact, for some 805 submittals, this number was a factor in deciding whether the quantitative criteria in Reg Guide 1.174 was met. For those who may not already know, which would be very few, if at all, that's the reg guide we use to make our risk-informed licensing action decisions.

The staff has significant challenges in establishing an appropriate approach to address this issue. After considering various relevant practical and operational issues pertaining to this problem, and giving due consideration to inputs that the licensees provided to us, we have been able to establish guidance on acceptable human reliability approach in this critical area for at least some parts of this problem. And we are working very hard to solve the remainder of it.

Next slide, please.

This example is rulemaking. Another example of importance to human reliability is the potential rulemaking relating to containment filtration strategies. To create the technical basis for this potential rule, we have to quantify the safety benefit of the filtered vents.

One critical input to this analysis is human reliability analysis. More specifically, the staff must use human reliability analysis to assign values for human error probabilities to establish mitigating strategies. Implementing most mitigating strategies involve activities conducted by humans outside of the control room. By the way, we also sometimes refer to them as flex strategies.

To that end, the probability of human errors, of actions performed by plant personnel outside of the control room, will influence the results of this analysis. Even though methods available to us to model human actions outside of the control room have not reached the same level of maturity as methods available to model actions inside the control room, we have a large number of tools and techniques to ensure qualitatively that these actions are feasible and reliable.

Now, I have to make another important point here based on some of the remarks you heard earlier with

respect to how we use qualitative and quantitative insights to make the best regulatory decisions. And I can do that because all our decisions are based on -- in addition to using risk-informed-type approaches, we use high quality inputs from what I call human factors engineering in combination with the numbers to make these decisions.

In that context, I would like to say with respect to mitigating strategies we capture both the principles of human factors engineering and the numbers from HRA to make the right decisions. For example, if you look at Section 18 of the standard review plan, and NUREG-0711, which is almost like my Bible on human factors engineering, it clearly articulates the fundamentals of human factors engineering that must be considered in developing feasible and reliable manual actions.

We have a plethora of other documents this agency has published to make sure that we can ensure safety and reliability and feasibility of these actions. For example, if you look at NUREG-1852, it delineates how we should assure reliability and feasibility of operator actions during fires.

And someone I think mentioned the flooding. If you look at the Appendix C of the Interim Staff

Guidance 12-05 it describes how we use qualitative factors to make sure that the actions that the licensee is relying on for external floods are feasible and reliable.

Now, after we make sure that those are feasible and reliable, at some point in time there is a necessity to do the best quantification we can. So we do that, too, because it is necessary for decisionmaking.

But one of the things I want to emphasize is I don't jump to the number. I have a lot of guidance out there to make sure that I do the right thing. Numbers are not my master, it's my slave.

Okay. What we do is once we make sure that the qualitative criteria are satisfied, we can then use well-informed judgment to assign failure probabilities for these operator actions. Due to relative lack of maturity of our tools in this area, we may have to place a heavy reliance on expert judgment in making sound regulatory decisions. Another point I want to emphasize is we can make good decisions today, but if you have better models we can make them more efficiently in a predictable manner.

In response to increases in future regulatory challenges in this area, we have a need to

increase our efficiency, clarity, and predictability by additional research to more accurately model these types of human actions.

Next slide, please.

Licensees also use HRA in a large number of applications. Actually, what I have to say here, the key message is -- Mr. Vaughn delivered -- but I still want to add one important point here. In addition to using human reliability analysis in areas such as licensees and oversight to engage the regulator, licensees, on their own initiative, use HRA to improve their plant safety.

They use it in design reviews. They use it in procedure updates. And also they use it in things like operator training, so that they can focus their operators to train on the human actions that are most risk-significant.

What happens is when a licensee does a PRA, they take the subset of the operator actions can be -- you know, that can be characterized as risk-significant, and they share that with the training people.

Now, most of you may not know, when they train the operators, the training -- the operators of the power plant, when you look at their training burden,

you kind of feel sorry for them because they spend like 20 percent time getting trained. So in the initial qualifications, you can't put in a lot. But based on my communications with -- and my personal experience by having worked as a licensee for 10 years, and recent communications with the SRAs, what they do is they use these insights into the requal where they have a lot of flexibility.

So the reason I say that, it's not mandated by regulation, but I think there's a powerful benefit to human reliability analysis that the licensees exploit, even though it's not required by the operators.

Next slide, please.

I'd like to conclude my presentation after making remarks on the relevance of the work that the Office of Research is performing in human reliability analysis. There are two important aspects to good human reliability analysis -- data and methods.

So let me first make the remarks on the methods. When I look for methods to do my quantifications for the operator actions inside the control room, I have a plethora of methods. So what I -- what could benefit me is something that would tell me the strengths and weaknesses of these different methods. Okay? And I think to that extent I want to

be thankful to Sean Peters and Office of Research for developing IDHEAS.

Now, I saw in a previous slide there was a statement that maybe the whole agency is not supportive of that. It may be a perception issue. We are using SPAR-H. IDHEAS is being developed. Okay? When we use what we are safe -- something we are safe with, when IDHEAS is ready, then we will go to that.

With respect to modeling complex human actions, those conducted outside of the control room, the situation is different. Our needs pertain to developing enhanced guidance to assist reliability of human actions outside of the control room. To that extent, a generic human reliability analysis methods supporting diverse applications that the Office of Research plans to develop will benefit us. Sean will give you details on that.

Next slide, please.

With respect to data, again, I am going to make two remarks; one with respect to data for inside the control room, and then on the outside of the control room. We tested to -- inside the control room for decades both the industry, and the licensees have been collecting data.

Now, more data help us reduce

uncertainties, enhance our clarity predictability. So that is useful to us. So because of that, I think we do appreciate the fact that Sean and his staff are working collaboratively with the plant to get more additional data from a simulator using a project called SACADA. I don't know why he named that SACADA, but that is what he called it.

So, on the contrary, when it comes to -- when it comes to collecting data for actions outside of the control room, that is an important area for us. That is, I think as Dr. Lyman pointed out, it is an area that we need to focus on getting more data on, and we have -- we got into communication with Office of Research to start that process.

And, in fact, what we are finding out is, as someone else said, one of the previous persons has said, there is data out there at the plants. The licensees have what they call job performance measures, all kinds of things happening. We have not -- we haven't started collecting that data in a manner that we can use it, but we have begun that dialogue with the Office of Research.

Now, that concludes my prepared remarks. My pleasure to introduce Sean Peters, Branch Chief, Human Factors and Reliability.

MR. PETERS: Okay. Thank you, Sunil. And I'd like to also thank the Commission for giving us this opportunity to present our HRA program. What I'd like to tell you about is what our HRA program does.

Our program in the Office of Research is -- for HRA, the purpose of that program is to build state-of-the-art HRA methods for the agency to use. We build good tools for our staff to use. Our needs are identified by both the user, by mainly Dr. Sunil Weerakkody's group, and by SRM. Three SRMs have helped guide our development activities over the last few years, and the primary one I am going to talk to you today about is the one listed first on this slide, which is the one on HRA methods. And it's SRM-M061020.

This SRM told the staff to -- or told the ACRS to work with the staff and external stakeholders to recommend a method or set of methods for the agency to use. My staff supported this activity by engaging -- and we saw this as an opportunity to -- as a developmental opportunity. Where we saw inherent weaknesses in many of the methods, we saw this as a way to improve those weaknesses. So I'm going to talk to you on the next slide about the activities that we have undertaken to address this SRM.

So one of the activities we first undertook

was an international benchmarking program where we took teams of international operators and ran them through simulated exercise at the Halden Reactor Project. We also simultaneously took teams of experts of HRA methodologies, and we used these experts to try to predict the performance of these crews at the South -- at the international -- or at the Halden Reactor Project.

This experiment brought up two questions. Number one, how applicable are these results to the U.S. crews? These are international crews at an international simulator. And also, when we ran this, we didn't get -- use multiple crews on -- or multiple analyst teams on one HRA method. We had -- basically each HRA team used one method.

So we saw a second benchmarking opportunity where we went forth, in collaboration with the South Texas Project Nuclear Operating Company and ran similar exercises at their simulator facilities. And we also took multiple teams using the same HRA method and we able to try to compare that analyst-to-analyst variability using HRA methods.

So the findings of the benchmarking study that -- the experienced teams, teams that were highly experienced, generally provided reasonable results

with their HRA methodologies. We found that -- and also, in addition to what we found in the best practices of HRA documents that we put out earlier, that all methods have particular strengths. They were all built for particular purposes, and they have strengths in some of those purposes.

But, then again, every method had a limitation here or there. So the other thing that we also found was that every method that we determined could use better guidance in one area or another of their methodology to help reduce some of that analyst variability.

Next slide, please.

So given the information we already knew by comparing HRA methods versus our best practices, and by the preliminary results of the U.S. and international benchmarking studies, we convened a workshop of international human reliability experts.

And the findings of the workshop -- we posed this SRM question to them. The finding of the workshop was that they didn't see one single method that was really suitable for all the NRC applications. They also saw this analyst-to-analyst variability as probably the single biggest issue we should try to tackle in our research programs.

So the staff got together and we came to the decision that we would use this as an opportunity to build an integrated method for the agency to use. We have done HRA for roughly three to four decades. We have developed the methodologies over those timeframes. And we wanted to take the pieces that we knew worked well and retain those, and we also wanted to improve on the areas that we know weren't working well, that the analysts pretty much had to work around throughout their methodologies. And we also wanted to maintain this focus on improving analyst-to-analyst variability.

Next slide, please.

So this is our third activity. It's the integrated method development. Basically, the goal of this development was to develop a methodology and reduce the variability and support the diversity of applications throughout the NRC. We wanted to conform to the ASME ANS PRA standard and the HRA good practices that we have developed over the years.

We wanted to retain those strengths that we developed and the methods over the 40 years, and we wanted to enhance key capabilities and key limitations in the state of practice, the ones that we could tackle. There are some that may be more challenging than others.

And we also -- one key piece we wanted to

have is we wanted to have a state-of-the-art scientific basis that was clearly linked to the methodology. And we also wanted this to be generic and flexible enough to support the diversity of applications at the NRC.

So next slide, please.

So what you'll see here is our strategic framework for method development. The top box here is the structured cognitive basis framework. This is our draft NUREG-2114 where we will be publishing it this year. ACRS has reviewed this, this scientific basis for human reliability, and the direct quote from the ACRS was that it contains valuable information to improve the understanding of the theoretical basis for human cognitive performance, the causes for human errors, and a structured framework to assess the contributions to error in the context of an evolving event scenario. It should be published, according to the ACRS.

We are publishing that. We see it as a -- at least a solid structure for our HRA-related work, and we are currently using it for some of our HRA and human factors-related activities in the Office of Research.

The next box down is our generic methodology for diverse applications. I'm going to talk a little bit about that here in a little bit, but

the final box is the IDHEAS method for internal at-power events. When the SRM was written back in 2006, the real issue at hand was that we had a plethora of methods for Level 1 internal events at power. And so this is where we started our work, but -- and we started working with industry. Industry is a key co-developer in the EPRI group, is a key co-developer of this methodology with us.

And we started down that path, but as you guys know, in 2011, all of a sudden we started having a more emphasized focus on events outside of the control room. We don't really have methods that were really designed for ex-control room activities. So we had to take this project a step back and realize, wait, we're starting to apply this into spaces like Level 2 and Level 3 PRA analysis. We are applying HRA in areas such as medical, as spent fuel storage and transportation, and as far as long-term waste disposal.

So these types of areas were areas we need to see we have a methodology that cannot just capture this Level 1 at-power, highly proceduralized control room actions, but we needed to have some framework to address these other domains, and the ones that are going to be more important to the agency in the future.

So we started this generic methodology

development in roughly that same timeframe. And the IDHEAS method, the ACRS also reviewed that, they identified some key enhancements, most of which we agree with, and we are working on those enhancements right now. And they also identified that they need a full scope testing of this methodology.

This full scope testing was also identified by our user offices and our users inside the agency, that before we roll this out we want this full scope testing. And this would be something that has never been done with an HRA methodology, to run through a full scope testing of it before use.

Next slide, please.

So we get the question, you know, how do we account for experienced operators? And how do they perform in these scenarios? Basically, each scenario has particular tasks that must be performed. And each of those tasks has certain demands, and those demands have associated performance-influencing factors. And you guys have -- we've talked a lot about performance-influencing factors, and the previous panelists.

We have like stress, we have distraction, we have fatigue, we have the design of the interfaces or the system, we have the training, we have the

procedures. These things compile together to create -- to basically -- they work against the cognitive limits of the operators. The operators can only think and cover so many details simultaneously.

Trained operators will handle these details much better. Trained operators with the right procedures will handle these details much better than people that aren't trained or maybe have lower quality procedures, and this leads to successes or failures. When you exceed those cognitive limits of the crews, you can lead to errors in that situation.

Next slide, please.

And so in the HRA process we go through -- we evaluate those PRA scenarios. It's a highly structured process and which I personally like to believe is more of like an expert judgment process. And we look at the scenario. Say this scenario we have a loss of reactor coolant pump seal cooling. We know the human action is that we have to trip the reactor coolant pumps to prevent seal damage and potential core damage down the path.

The reactor operators can either trip the reactor or they don't. If they don't, you can lead down a path towards failure. And if they do, you have a path towards success. And we tackle that particular

scenario through both the qualitative analysis, which what I would say 90 percent of -- 99 percent of all HRA practitioners view as the most important piece of HRA is this qualitative analysis process, where you try to really understand the scenario and you try to identify those human failure events that are associated with the PRA, and you try to analyze the tasks that the operators have to perform.

And then you go down through the human failure quantification, where we identify, okay, now that we know the tasks you have to perform, how can they fail at these tasks. We analyze those performance-influencing factors that we get from that contextual information from the event, and then we go through this expert process of estimating the human error probability associated with it.

As I want to just restate, most HRA practitioners view that last step, the analysis of the human error probability estimate, as the least important of all the steps. We don't gain the insights necessary from the human error probability but from the rest of the structure of the accident progression.

Next slide, please.

So our basis for our IDHEAS methodology is that our humans, our teams, they perform their tasks

through these cognitive functions. As on our previous example, we needed to trip the reactor coolant pumps, and we have these various underlying cognitive functions. We have detection, understanding, decisionmaking, and action. So in this particular example, when you detect an alarm is going off, when you check those plant parameters and see what the actual problem is, that sort of detection stage, we need to understand what the plant is doing, and we need to make the diagnosis steps of diagnosing that we lost that seal cooling.

We also need to make the decision, oh, great, we need to trip the reactor coolant pumps. And, finally, we need to execute our procedures to actually do that trip of the pumps.

Next slide, please.

And so what you'll see on this slide is a structure of how these particular cognitive functions are laid out, I think explicitly linked in our IDHEAS methodology. That we have that task, that human event, which is the reactor coolant pump seal cooling loss. We have to monitor the plant, diagnose that problem, follow our procedures for taking care of that problem. And we do that through these cognitive functions.

And potential failure modes for that is,

say, one, we did not attend to the alarm, say we didn't see it, or we didn't understand the alarm or the data that was presented by the plant. Or we delayed our implementation because we had numerous competing priorities that may have taken precedence over this particular area.

And there are various performance-influencing factors that can go into that, and we can be distracted. May we have command and control issues, we have alarm design, we have a perceived urgency of other tasks that prioritize over this one, or we even may have procedural or training issues as we've seen in other events.

Next slide, please.

And so then we -- as a last step in the process, we estimate our human error probabilities, where these error probabilities vary based upon the complexity and the combinations of the influencing factors. And these failure scenarios were estimated through a formal process, expert elicitation process, using experts in operation, human reliability analysis, PRA, and the cognitive sciences.

And what you'll show -- and this is how the IDHEAS methodology is set up -- that the more complex tasks have a more likelihood for failure, just as we have

seen in real-world events. And a simpler task with fewer negative performance-influencing factors have a higher probability for success.

Next slide, please.

So the ACRS has looked at the IDHEAS methodology and has reached the conclusion that there are key elements that will reduce the interanalyst variability. And these particular improvements that we have made with the IDHEAS methodology include that we have taken the bits and pieces of various HRA methodologies and taken those strong pieces.

We provided guidance on every step of the HRA process. Many methodologies don't have guidance on all steps or what I would say complete guidance on all the steps. We have enhanced guidance on the qualitative analysis and task analysis. We have seen particularly where these areas being what we consider the most important part of the HRA, these are areas that we have enhanced that guidance.

We have an explicit model for the human cognition, and we have linked that to the human failure modes. And we have explicit guidance for performance-influencing factors. The basis for those performance-influencing factors linked back through our scientific literature for human factors, and we have

questionnaires to help the analysts assess those performance-influencing factors and the assessment of the methodology.

And, finally, we estimated these human error probabilities through expert panels as we don't have significant enough data to do it analytically yet.

Next slide, please.

So initial testing of this methodology, three HRA analyst groups independently tested the IDHEAS method. And I call this preliminary testing because it was just a proof of concept of the methodology. We found that the parts worked as intended. There are key -- we believe key improvements to the limits in the state of the practice. There is good traceability, clear documentation.

We have what we consider, based upon our three results -- take it for what it is -- that we have some reasonable interanalyst variability. There is more analysis effort up front. There are simplified methods that don't pay good homage to really understanding the scenario, so there will be more interanalyst -- or more analysis effort up front for understanding that scenario and laying it out.

But we believe it reduces deliberation on the back end. This is what they have seen on the -- when

you have one analyst team create their model, say the industry does, and we create our model, and then you argue about those differences in the model, we reduce that because we have the clear pass to run through the methodology.

And we also -- given that this is a 300-plus page document, it is not as easy for the users to use as they would like. So they desire clear user-friendly implementation guidance, so we are working on developing a user's manual for the users to promote that.

Next slide, please.

So we also -- given the context that we -- we are also developing a generic methodology. Given the fact that we have, you know, other areas, Level 2/3 PRA, reactor shutdown operations, external events, fuels material byproduct applications of all interest to the NRC, we have had to develop -- we have had to think in the larger context than just this Level 1 model. And so we have been-- we have begun developing that. And this will allow us to tackle a broad spectrum of human actions, including ones without detailed procedures or ones performed by people outside of the control room or non-trained operators.

We may have complicated decisionmaking, which comes from the technical support center or

operational support center. We also have the performance-influencing factors that we don't typically experience or use in a control room, like you are in a high radiation field or you have floods or you have fires. These kind of things aren't typically assumed in the current methodologies.

Next slide, please.

So the path forward. The cognitive basis framework, we are publishing that this year. We are using it in the NRC's human factors and HRA engineering. Our IDHEAS methodology, we are going to take that, make many of the enhancements that the ACRS is recommending, and we are going to be testing that for NRC applications with our users and with industry. And we are developing this generic methodology. We are currently tailoring it for the containment filtration strategies rulemaking and using some of the insights from that methodology for that rulemaking effort.

And we are also working with NRR to help guide us towards the other areas that we want to tailor that generic methodology for, and we'll be finalizing our user's guidance in the 2016 to 2017 timeframe.

Next slide, please.

So I wanted to just briefly talk about the other activities that help inform this methodology, and

we have an expert judgment guidance SRM that told us to develop the standardized expert judgment method for the agency to use. We are heavily using that expert judgment guidance insight in both the IDHEAS methodology and in the Level 2 and 3 PRA that the staff is performing. And we also have a very well-developed HRA data program where we developed the SACADA database, which Sunil referred to earlier, where we are working with the South Texas Project Nuclear Operating Company, collecting all of their live simulator training data, and we are also collaborating with international partners like Halden and other -- and a couple of other countries to collect their data and share data on the human performance and simulator scenarios.

We are developing baseline human performance data at our university partners. We have -- the NRC or my group owns two pressurized water reactor simulators, one that we have at the University of Central Florida which helps us tackle some of the questions about these performance-influencing factors, and we also team up with the Halden Reactor Project to do what we consider high fidelity experiments with operational crews to test some of those insights that we gained from University of Central Florida.

So next slide, please.

So these data sources, these are -- this is just a picture of the various data sources. The top left is our team of human factors operations, cognitive sciences, and HRA practitioners that helped develop the SACADA database at the South Texas Project. The top right is our NRC-owned human performance test facility at the University of Central Florida. And the bottom picture is one you guys have probably seen a ton of times, which is the Halden Reactor Project, where we do all of these targeted human performance experiments.

And our concept with this data is that we will take the data that we have, and we will try to validate the HRA methods that we -- that our IDHEAS methodology used. So there are explicit linkages and very similar structure to the SACADA database to the IDHEAS method-based sciences. And so our concept over the next few years is to try to prove this concept of using that data to back-inform the human error probabilities and the IDHEAS methodology.

And I'd like to pass my presentation over to Rich Correia for the conclusions.

MR. CORREIA: Thank you, Sean. In closing, Commissioners, human reliability analysis is used to support our regulatory activities. The staff developed the IDHEAS method, as you have heard from

Sean, which was done in collaboration with EPRI and other stakeholders. We are also developing a generic method that can be used -- tailored for multiple applications. These methods were developed using state-of-the-art technical analysis and operating experience.

Finally, as part of our human reliability analysis program, we will seek to improve our methods, and we continue to test them and collect and use more human performance data.

This concludes our presentation. Thank you for the opportunity.

MR. WEBER: I would just add in closing that appreciate the close collaboration among the offices, particularly the Office of Nuclear Regulatory Research, the Office of Nuclear Reactor Regulation, and the Office of New Reactors. While you didn't hear a presentation about the application of these methods in other areas, we are also, as we are developing this generic methodology, thinking about how would we apply this to broader apply across the responsibilities of the agency.

Thank you.

CHAIRMAN MACFARLANE: Great. Thanks very much, guys.

Commissioner Apostolakis.

COMMISSIONER APOSTOLAKIS: Thank you. First of all, I like what Mr. Weerakkody said about the integrated approach. But you did start with a model that has no justification, the SPAR-H, which brings me to what I said earlier this morning.

You really have to develop application-specific guidance, not as a side project but a major effort should be there. Developing a generic methodology is okay, but, for example, the significance determination process, can you develop guidance just for that, taking only what is appropriate from the generic methodology and give step-by-step guidance?

The flex methodology, we have asked the industry to tell us or to explore what can go wrong. I mean, transporting this heavy equipment under extreme conditions is not a straightforward manner. So having a generic methodology is good, but it would be better if you had specific guidance how the staff would evaluate the feasibility and reliability of these actions.

Sean, you mentioned that you are already doing something about the filter vent strategies, which is good. So I -- that's what I have in mind. I mean,

here is what you will do, and this is a simple model, but it is based on a more sophisticated model.

In my view, this is why ATHEANA did not catch on. It was too elaborate. And speaking of ATHEANA, have you guys explored why a model that was advertised as a great model 10 years ago now we don't even talk about it? Are there any lessons learned there other than it was too complex for the average user? Where is ATHEANA now?

MR. PETERS: Where is -- okay. I completely agree with your statements about making simplified methods model-specific. And on the strategic framework slide, we show that the generic methodology is just something to be all-encompassing, that it's a standard framework for HRA. So when we have these particular items like SDP, we know we need to sub-select it to make it more useable. So I completely agree with what you're saying.

This is kind of the concept that we have right now, that for each method, for each detailed use, we will have this standard scientific framework and we will then build simplified methods based upon that standard scientific framework. So completely agree.

And we haven't really done a lessons learned with the ATHEANA methodology, but the feedback we have

constantly gotten was the difficulty of use, the amount of effort and resources that are put into that. It is not completely dead, however. It is a piece of the fire HRA. It is a piece that some industry participants are using to help understand that qualitative analysis piece of the fire HRA.

And so from that standpoint, ATHEANA had some very great pieces and qualitative analysis that we are trying to capture into our IDHEAS methodology. And even pieces that were recommended to us by the ACRS to incorporate into our IDHEAS methodology. So I can't say that the IDHEA concept of ATHEANA is totally dead, but it is living on not just in fire HRA but in our IDHEAS methodology.

COMMISSIONER APOSTOLAKIS: So there are still elements of ATHEANA that are useful to what you are doing now, as it would be expected.

I remember a very interesting simulation exercise at Halden where they had -- where finally they are using U.S. troops -- I mean, crews, right? Because a question the ACRS asked a long time ago, what are the results of simulation exercises that use Swedish crews and a Norwegian operator, what are they telling us about American operators? Well, you took care of that.

But in one particular exercise, I remember

several crews responded to a particular accident sequence in five to six seconds. And one crew took 11, 12 seconds. It was clearly an outlier. And I've been wondering, what does that tell us? I mean, is that something that is included in IDHEAS or in other models?

I know that we are using the time available and time required in the fire analysis. Do we do that in other applications as well? I mean, one crew was completely off.

MR. PETERS: Sure.

COMMISSIONER APOSTOLAKIS: Doubled the time to realize what is going on.

MR. PETERS: Yes. What I've seen -- and I may get over my head very quickly, so I'll rely on Dr. Xing to correct anything that I say that's incorrect, but what I have seen is that we do, in fire HRA, and we do have that piece of time available versus time required, that we are retaining into the IDHEAS methodology.

The ACRS has recommended a slightly different approach in their letter last week, and we are looking at various alternatives at this moment. So I'm not going to say for definite we are keeping this time available versus time required methodology that is there, but that is one piece we're looking into.

And how they've addressed that over time is -- what I remember is that if you have double the time that you really think it's going to take, that from that point it's a feasible action. If it's less than double the time, it's a non-feasible action, and that kind of two-step function of granularity may not necessarily be useful to the analyst as --

COMMISSIONER APOSTOLAKIS: But when you are eliciting expert judgment to quantify, the experts are not thinking in terms of time, are they?

MR. PETERS: They do assess the feasibility of the action and the time required to do those actions. It is one of the considerations in that process.

COMMISSIONER APOSTOLAKIS: As a final comment, I read the ACRS letter and it seems to me the Committee was trying to compete with you in the number of pages.

(Laughter.)

It was an incredibly long letter.

MR. WEBER: Very thorough.

COMMISSIONER APOSTOLAKIS: Sure. But it was long. Thank you very much.

CHAIRMAN MACFARLANE: Commissioner Magwood.

COMMISSIONER MAGWOOD: Thank you, Chairman. The previous panel discussion with the Commission, it sort of highlighted a variety of cases where, you know, people can either do something very positive and beyond procedures, or they can fail to implement procedures for one reason or another.

So, and I wonder, in thinking about that, and there was discussion about the uncertainty that goes with this and the other panel discussed this, but in thinking about it, you know, in a very crass way, HRA is an attempt to reduce the individual to the same type of functionality as a pump or a valve in a PRA.

And pumps and valves in PRA have extraordinary behaviors, too. You know, there were pumps that lasted far longer than they were supposed to, and that is extraordinary and it's a good thing when it happens. But you don't count on it, but so it's -- you have pumps that, for whatever reason, just fail very quickly, and, I mean, there is no clear explanation for why that particular pump failed. And that's something that -- you know, that to some crass way people are sort of like that, too.

First, let me -- I have a followup on that, but I don't know if you want to comment on that. Is that --

MR. PETERS: Yes. I think my concept with HRA is that we do try to model an average for, say, someone that we consider the middle of the road, because we are trying to get these probabilistic insights. So, yes, you will -- for all these cases where you have a heroic action, you have somebody running forth and taking charge, you have another guy who is running the other direction. And we have examples of that in Fukushima, and we have examples in other major catastrophes.

And so from that standpoint, those are hard to capture when you're doing a probabilistic assessment. You're trying to say, what are these probabilities? So for us, for those insights, it's best gained when you are trying to do a -- like a predictive HRA. You are doing something that is more middle of the road or more average. But when you're doing a retrospective analysis, and you can actually see what people did, you can then take HRA and say, "Okay. Now I have this many people running away, and this many people running into the fire. Now what are the possibilities for success or failure based upon that knowledge?" So we can look at it both ways.

COMMISSIONER MAGWOOD: And I think that I'm sort of conceptually -- and I recognize people aren't

pumps and valves, but conceptually it's very similar to the way we -- that we analyze pumps and valves and just -- so there are extraordinary things that happen on both sides of that, and you try to take an average. And that's just the basis of the analysis. And as long as you understand the limitations of that, you can apply this as a tool.

I think you mentioned that you anticipate a full scope test of IDHEAS at some point. Can you elaborate a bit more on what that means?

MR. PETERS: Well, yes, I just read the -- like a rough draft of the testing plan yesterday. And I'd actually prefer the person who wrote the plan to answer that question. So I'll pass it to Dr. Xing.

COMMISSIONER MAGWOOD: You've been trying to get her up here all day.

MR. PETERS: I've been trying to get her up there forever, yes.

DR. XING: Hi, Commissioners, and ladies and gentlemen. I am Jing Xing. I'm the technical leader for developing the IDHEAS and this whole suite, the project.

As far as the full scope, there really isn't a very scientific definition. So, but the minimum criteria we would like for testing will be we should

prove the methods are working, and we should demonstrate the delta between this method and our current practice. And we should demonstrate it's easy to use for people, that's our testing goal.

And this testing goal we develop has the basic requirements of what we need. For example, we need to get our users fully involved for what they expect for testing. Then, to determine the scope. So for the large -- at a high level, the scope for the testing scope, we should test this method. To cover -- to use the -- I would say a good enough number of testing teams, because we want tests, variety of the analysts, the test -- it's between the analyst team. And we want to use the scenario that covers from easy to difficulty, and also cover our current application that we should include in SDP scenarios, and the challenges used in simulators.

So, and also, we need to develop for the user acceptable testing criteria. Testing for what? So we want to test its accuracy. Maybe for some rare event you can never be able to reach the accuracy. So what is a good enough criteria?

Those are the things that -- in our testing plan. And it's not an ideal package, but I would say it's -- as Sean said, this is -- it's the most

comprehensive testing for the HRA methods that have developed so far.

MR. WEBER: Okay. It sounds like we are still developing it. And we'll, I'm sure, work with Research and NRR to ensure that as the test plan is formulated that it is responsive to our regulatory needs.

MR. PETERS: Literally, this is a one-week-old process at the moment. So --

COMMISSIONER MAGWOOD: I understand. Thank you.

I think Commissioner Apostolakis mentioned, just in passing, how you are applying HRA in the -- to the -- in the significance determination process. Is there clear guidance on how to do that? Is that something that the staff is using routinely now?

DR. WEERAKKODY: Yes. Let me answer this. What we have done is, especially with significance determination process, we have to recognize it is a process where we have made timely decisionmaking, but we had to make good decisions.

So what we have done is we have created another guidance, what we call -- we call it fast guidance, risk assessment. I can't remember SNP stands for. But what we do there is we identify some critical

areas in HRA that might be what I call pinch points. So there is a separate guidance for that.

COMMISSIONER MAGWOOD: Okay.

DR. WEERAKKODY: Yes.

COMMISSIONER MAGWOOD: Let's see, I think that's the end of my questions. A couple of quick comments. First, I just wanted to thank the staff for working so hard to collaborate. We have a lot of partners here who are engaged in this work, and it's very satisfying to see that we have not been insular in this. We have reached out quite broadly to a wide range, and I think that has been very productive and very beneficial.

I also wanted to gratuitously recall one of my old professors, which -- who would be very amused by this conversation. He was actually a philosopher, and his specialty was -- many papers he wrote over many decades about how the human mind can know what the right thing to do is but yet still do the wrong thing. And to some degree, we are kind of having that conversation, not in the philosophical sense, but in a real sense.

So I wish he were here to sort of sit and listen to this conversation. I'm sure he would opine about this.

Finally, I wanted to end with a question or

a clarification from Commissioner Ostendorff. I believe I heard him make -- use the phrase "shoot a torpedo." And I was always under the impression that one launched a torpedo.

(Laughter.)

And I was hoping that as I hand the microphone over to Commissioner Ostendorff that he could provide some clarification on that terminology.

Thank you, Chairman.

CHAIRMAN MACFARLANE: Thank you.

Commissioner Ostendorff, maybe you can provide some clarification.

COMMISSIONER OSTENDORFF: This is a very sensitive issue.

(Laughter.)

The correct -- the equipment used on a torpedo -- on a submarine to launch a torpedo is called the Launch Delivery System. But for the cowboys, for the people in the club, shooting a torpedo is typically the terminology used on the boat. But thank you for paying attention.

(Laughter.)

Thank you all for your presentations. I particularly appreciate that Sunil and Sean used specific examples, which I found very, very helpful.

But I'm going to kind of bore down on one of the examples of Sunil here, because I want to make sure that I understand where you're headed.

I'm looking at your Slide 9 on containment filtration strategies. And a couple of comments that I understand -- we should look at, first, the need -- and I agree there is personnel actions outside of the control room that come into play here. Do I understand, though, that you are trying to assign a numerical probability of success of those operator actions outside the control room from the standpoint of the rulemaking or filtration strategies?

DR. WEERAKKODY: Yes, Commissioner. Yes.

COMMISSIONER OSTENDORFF: Okay. That causes me some concern perhaps, because I've heard -- I think I heard both you and Sean say that there is no generally accepted method.

DR. WEERAKKODY: If I may elaborate, I think --

COMMISSIONER OSTENDORFF: Please do, because I want to get --

DR. WEERAKKODY: Yes. I gave you a short answer. She said yes/no answer. But to kind of get the context, you have to understand the process you go through. First off, we ask large number of questions.

We will be asking large number of questions from the licensee about their procedures and the guidance and the training, whether they'll work during the environmental conditions that will be present during the accident. So there's that qualitative piece.

We look at that as a first step to make sure that we have good actions that are feasible. Okay. So that's the first thing we would cross.

Now, the second step is, at some point in time when you try to do the cost-benefit analysis, you -- the next step is you assign some screening numbers. Now, when I say "screening numbers," you might say for activities outside of the control room there's a 30 -- there's a 70 percent chance of success.

Now, I may not have a very, very scientific basis for that, but we have a lot of experience with respect to looking at different procedures and coming up with a reasonable screening number. For example, we know it's not going to be one in a hundred, because it's done in -- outside of the control room by people and there's a lot of challenges.

Now, after we do the screening analysis, that can highlight some of the key things that we need to fully explore. Now, my knowledge with respect to how we would go from there to doing the actual cost-benefit

analysis stopped because my --

COMMISSIONER OSTENDORFF: Whoa, whoa, whoa, whoa, whoa. Where do you get the cost-benefit analysis? I thought we were talking about human -- the probability of human action -- humans performing acquired operator actions.

DR. WEERAKKODY: But to do the cost-benefit analysis, you have to come up with, what is the safety benefit of this particular proposed change?

DR. UHLE: Sunil, can I help out a bit?

DR. WEERAKKODY: Yes, please.

DR. UHLE: Hi. My name is Jennifer --

COMMISSIONER OSTENDORFF: I want to -- I've got limited time. Jennifer, I'm fine with you being there. I want to focus, though, on -- my key -- I'm not interested in the cost-benefit analysis. I want to understand, though, how you are quantitatively assessing the ability of an operator outside the control room to perform actions associated with containment venting or filtering. That's what --

DR. UHLE: Okay.

COMMISSIONER OSTENDORFF: I've got limited time here.

DR. UHLE: Again, with our limited

analyses or methods that have been benchmarked for complex scenarios outside of the control room, we really look at human factor insights. So, for instance, where is the equipment? Is it easily retrievable? What would be the operating conditions?

COMMISSIONER OSTENDORFF: Those are your Part 1, which I agree with.

DR. UHLE: Okay.

COMMISSIONER OSTENDORFF: I'm fine with -- I'm concerned about Part 2.

DR. UHLE: Okay. And so we have, you know, this -- this I would say qualitative view. When we go to do the technical basis for the rulemaking, we have to understand the benefit, the safety benefit of this action. So we are arranging the values, we are doing sensitivity studies. If it turns out that the technical basis highly depends on these numbers, then we'll be diving in deeper and perhaps -- well, we are going to plants to see these actions taking --

COMMISSIONER OSTENDORFF: I'm sorry, but I've got limited time. I don't think you're answering my question. I think the question I'm getting to is I understood Sunil as saying that you're going to assign some quantitative number --

DR. WEERAKKODY: Yes.

COMMISSIONER OSTENDORFF: -- to the likelihood of a particular operator action being completed outside of the control room as part of the containment filtering strategies procedures for a particular plant.

DR. WEERAKKODY: Correct. Yes.

COMMISSIONER OSTENDORFF: Isn't that -- that's what concerns me.

DR. UHLE: Right. But we're using -- we're doing that in a range of values. We recognize it's not a precise value, so we're doing several sensitivity studies. And, for instance, from the human factors approach, if it's highly likely, okay, maybe that's 70 percent. If it's -- you know, if it's moderately likely, maybe that's 50. If it's not -- if it's very unlikely, then maybe that --

COMMISSIONER OSTENDORFF: Okay.

DR. UHLE: -- is 10 percent.

COMMISSIONER OSTENDORFF: That concerns me, just as an individual Commissioner. I'm not -- I'm just telling you that I'm trying to understand because other comments that you made and Sean made about the lack of -- you know, lack of agreed-upon methodologies, when you're trying to quantify something that we perhaps don't necessarily have agreed-upon consensus tools yet, can

be quantified.

DR. WEERAKKODY: Can I say something?

DR. UHLE: Sure.

DR. WEERAKKODY: Yes. I think the part that we did not mention is --

COMMISSIONER OSTENDORFF: Thank you, Jennifer.

DR. WEERAKKODY: -- again, if you are looking for, yes, here is an absolute number, it's scientifically 100 percent correct, we are not there, but we deal with the sensitivities. But what we do look at is close look at the operating procedures, some of the other procedures they look at that -- that they will be using. And we have lot of experience, Commissioner, in terms of assigning consensus-type numbers to similar kinds of situations, number of other applications. So --

COMMISSIONER OSTENDORFF: Okay. So I'm going to make a comment here, and then I'm going to ask you to respond to it, because, again, I'm watching the clock here. I know this has already been a long meeting. But I've got to tell you, you know, the Chairman raised comments about uncertainty earlier. Commissioner Svinicki raised -- which I agree with. Commissioner Svinicki raised comments about

extraordinary actions by people under difficult circumstances.

I fought fires on submarines before, and I would laugh at anybody trying to model the ability of somebody numerically to successfully fight a fire. I've done it before, and I would -- I'm sorry, I think the credibility factor there is really key to me.

Commissioner Magwood made a comment, previous questions about concerns on equating pumps and valves with people, which I agree with, that -- so, you know, I go back to Jim from the previous panel talking about integrated crew response, where there's backup of other people, if somebody makes a mistake, where is somebody else going to weigh in to back them up? Those are things that are very difficult to assign a number to, yet those are real operator actions and real responses.

So I'm a little bit skeptical of what I saw on the note page for this slide, because I'm hearing some inconsistent things from our staff about, well, no, we don't necessarily have good models to agree to, yet you're going to try to use these on a rulemaking that is very important to the Commission in the near term.

DR. WEERAKKODY: If I --

COMMISSIONER OSTENDORFF: I'll stop there.

So, please, I've said a lot.

DR. WEERAKKODY: May I say, I think -- I don't want to leave a concern with you, because I think one of the things we did not mention is I'm not going to argue with, you know, how uncertainty, you know, the numbers there, but every regulatory decision we make gets risk-informed. In other words, this -- whatever the number we come up with is, one of the four criteria we look at, whether it is this or SDP, we are looking at things like defense in depth, safety margin, so when we make a proposal on anything, or we make a decision, we do give a hard look at those --

COMMISSIONER OSTENDORFF: Yes, but you're -- I'm talking specifically about the operator action piece and your statement that somewhat alarms me. And I highlight that because this rulemaking is before -- you know, is something that is very important to the Commission. I certainly agree with your first step. Are these actions feasible? Can they be done? Can you observe these in a simulator in the plant? And so I think I completely agree with that Part 1.

But that Part 2 piece of trying to assign a number, man, I will tell you from experience in the military and in the nuclear plant operations military, also through some family experience in combat recently,

that the military doesn't try to assign a .63 percent that this soldier is going to shoot that insurgent without having to get backup from this person over there. That's just -- you train and practice and you identify those errors, and you try to reduce those errors to as low a level as possible through training and repetition. But trying to have a regulatory basis rely upon numbers the way I'm hearing you talking about it -- and, Doctor Uhle, I have a little bit of maybe some healthy skepticism at this point. So I will leave it at that.

Thank you. Thank you, Chairman.

CHAIRMAN MACFARLANE: Amen.

All right. First, I'll start off with a question, and then I'll continue on that line of interrogation. Do regulators in other countries use human reliability analysis? And, if so, which countries, and how do they do it? How do they use it? And you can take that for the record if you want.

MR. PETERS: Yes. I mean, there are regulators that use human reliability analysis for their regulatory decisionmaking. The ones I know of are, as Claire had mentioned, that they were using it in the United Kingdom. They use it in France.

CHAIRMAN MACFARLANE: How do they use it in

France?

MR. PETERS: Well, this is an area beyond my knowledge. I'd like to pass it to one of our international experts to talk about that.

CHAIRMAN MACFARLANE: All right. I don't want to spend too much time on this, but just real brief.

DR. TAYLOR: That's okay, because I don't have a lot to say about it.

(Laughter.)

CHAIRMAN MACFARLANE: Okay.

DR. TAYLOR: My experience is only from the UK. The UK ONR, the Office for Nuclear Regulation, they -- to the best of my knowledge, they don't actually do their own HRA, but they very -- they review the HRA provided by the licensees in great detail. And they also have the possibility to comment on that and ask for additional analysis where they see fit.

CHAIRMAN MACFARLANE: Okay.

DR. TAYLOR: So that's how they use it. I don't have any insight into how they use it to actually make regulatory decisions, just in terms of their review of safety cases.

CHAIRMAN MACFARLANE: All right. Well, that's helpful. It would be helpful to know how it's used in other countries. You knew that question was

coming.

Okay. Back to the filtration rulemaking. So how are you characterizing and calculating the uncertainties?

DR. WEERAKKODY: I want to be -- make a distinction between when you say uncertainties, you know, there is a big parameter concerning it. In other words, we have tools if I wanted to say mean is .1 and then I want to throw in a distribution and calculate that. But in this particular case, Chairman, I think what we would rely more on is in the sensitivity. In other words, we would say, okay, for this operator action, my screening value is 30 percent or .3.

Then, I might say, well, what if I am wrong? You know, let me try and get the 4.5. How would the decision be impacted with that number? So that's how we would deal with the potential answer using that number, using sensitivities for this particular case.

And then, of course, when it comes to the older edition, we rely on the other factors of risk-informed decisions, what does this do to defense in depth, safety margin, and so on and so forth.

CHAIRMAN MACFARLANE: And so for this analysis, where are you getting your input data from?

DR. WEERAKKODY: Again, I think you are --

let me answer it in a general way. I think that's a general question for every HRA in terms of --

CHAIRMAN MACFARLANE: Yes, sure it is.

DR. WEERAKKODY: Simulators.

CHAIRMAN MACFARLANE: Ah.

DR. WEERAKKODY: Okay?

CHAIRMAN MACFARLANE: Okay. That goes to the next question, and, you know, Sean, you showed your data sources. Your data sources are all simulators, which are models. And models are not data. They are models of -- they are models of reality. They are not reality. So this goes back to what Commissioner Ostendorff was saying where you -- actual experience is very different, you know, from a simulator.

So if you are informing your models with model results, what does that mean? Is that meaningful?

MR. PETERS: This is not the only source of data that we have. It goes into the SACADA database. We are actually modeling actual events that have taken place. So the H.B. Robinson event that Sunil was using earlier, this is one of the first pieces that we're putting into the SACADA database.

We have also been working for 10 years prior to this modeling all of the augmented inspection team

and IIT events that have come through the agency through our previous database that we had, the HERA database. So we have all of those events already modeled in our previous database, and we are actively moving those models on the SACADA database.

So we have actual events that we're putting into that. And when we're looking at the psychological underpinnings of our IDHEAS methodology, we have actual scientific data on the various performance shaping factors like fatigue. So you run these people through events, and they experience fatigue, how they perform or how they don't. And so we have lots of data when it comes down to these individual factors that we've identified through psychological testing.

DR. WEERAKKODY: Thank you, Sean.

MR. PETERS: Yes.

CHAIRMAN MACFARLANE: I think this is where you -- you know, yes, you have discrete data points on fatigue or, I don't know, confusion or whatever, but this is where the interactions of these different situations are incredibly important. And I think we all know from our own experiences in life that trying to really make predictions about how we might behave, or how others we know well might behave, it's really difficult.

MR. PETERS: It is difficult, and some predictions are much easier than others, like, say, you have a stop sign. Hey, 99 percent of the time people are going to stop at a stop sign, or at least do that roll through. That other one percent, well, that's a whole different can of worms.

CHAIRMAN MACFARLANE: Maybe for you. So let me ask another question. In the previous panel, there was a lot of discussion about this, too, the use of expert judgment. Okay? So you use expert judgment. You mentioned who some of your experts are or vaguely, general categories.

How do you evaluate the quality of this expert judgment? How do you characterize the uncertainties associated with this expert judgment? Isn't expert judgment simply opinion dressed up in pretty clothes?

MR. WEBER: Well, Chairman, the Commission has tasked the staff with developing guidance on how to use expert judgment. So, Sean, that's in your group.

MR. PETERS: Yes. And Jing is our lead, and she would like to -- she is chomping at the bit to answer this question.

DR. XING: Okay. I have been also the technical lead for developing the guidance for expert

elicitation, expert judgment. So we developed our initial work package to recommend the agency to use based on -- or start the process as it has been exercised many times, the SSHAC process. And we exercised that process in our IDHEAS expert panel.

So it's a structured scientific process, and the very first step of the process is to establish good data and knowledge base. In that process, we try to collect all kind of data, not just from simulator, but, as you two already mentioned, from other domains -- aviation and the manufacturing industry.

And, fortunately, IDHEAS, because it's based on this cumulative basis we developed, it naturally has a way -- allow us to judge how we can use those data in the other domain, whether like compared to the Air France accident you just mentioned. So we know people are still doing the same kind of work, a combination decisionmaking. And what factors are different, how that would impact a similar situation in nuclear power plant.

And also, the expert judgment process, tried to maximally fully elicit those uncertainty factors around every topic with -- see, we are judging not just the probability of this failure, but we have the different group of people think about from

cumulative or social science aspect what factor will come in, other factors can make this fail, what are the individual performance differences.

And also, from the operator side, provide us knowledge as far as operation, we have all kind of mitigation strategies. So we take all of this into consideration.

CHAIRMAN MACFARLANE: Right.

DR. XING: And build a distribution of probability.

CHAIRMAN MACFARLANE: I just worry that we have actually qualitative information, which is fine. I'm fine with qualitative information. I think we should examine a lot of these factors. But we shouldn't then all of a sudden pretend that it is quantitative, just assign a number to it and then use it in a calculation that produces a number that's really meaningless because it was qualitative to begin with. So that's a concern.

One more quick question for Sean. You talked about experienced operators. Do you always assume an experienced operator? What is an experienced -- what's the definition of an experienced operator?

MR. PETERS: I'm not sure of a formal definition of an experienced operator, obviously

somebody who has been doing it for a number of years. For our expertise, or for our experience, we found that people that are really in operations training have the most insights into human performance, because they get to see a litany of crews run through experiment after experiment. And they will see the relative level of failures of those particular crews.

So basically, the people that have been doing it for numerous years, in their particular context --

CHAIRMAN MACFARLANE: Experience is valuable. Certainly, I want to go to a surgeon who has done a lot of the same surgery and not one who is new at it. But, still, I -- there's still, you know, that qualitative element there.

MR. PETERS: Oh, yes. There is.

CHAIRMAN MACFARLANE: Okay. I will stop and turn it over to Commissioner Svinicki.

COMMISSIONER SVINICKI: I'm kind of smiling to myself, because as I prepared for this meeting I thought might be the least enthusiastic person about HRA. But I'm beginning to feel like one of the greatest defenders of HRA, or maybe I'm just a contrarian, so I'm becoming a defender of HRA.

You know, the reason I asked my question

about looking only towards human error and not towards human superior performance is I made the point that, you know, any methodology we're using we should want to have the ability to compare that against real-world results. If you say, well, as Commissioner Magwood -- I'm sorry, but I thought I heard you say, if a pump runs longer, that's great, but you can't count on it. But I think, you know, I don't want to count on everyone being superhuman, but on the one hand it's not real-world results if no one is superhuman.

So I'm just trying to, in my ignorance, get some sense, and I think the answer I got was that, at least for nuclear power applications, there is no balancing of factors on the positive side. So it was just -- I just wanted an awareness of what some of the limitations of the model are.

But, you know, the record will reflect that on complex rulemaking packages my vote is frequently the last to come in, and I know that might be a source of frustration at times. But one of the causes of that is spending time with the reg analysis, the tech analysis, and other things that aren't, you know, in the strictest sense things that the Commission is voting on, but they are the underlying analytical work that was done that takes -- and that's why the rule language the staff

proposes, it looks like it looks.

Also, for cost-benefit -- and I'm a rather substantial proponent of cost-benefit, and I think regulation should have a benefit that justifies their cost. So when I look at the staff's analysis, I see that you have to use all kinds of subject matter experts, expert elicitation, and I want to compliment Commissioner Apostolakis. This hasn't been acknowledged today, but one of his early focus areas when he came on this Commission -- it might have been his first COM -- was on expert elicitation and having some sort of extremely scrutable and consistent approach to its use. And at the time that he wrote that COM, I'm not even sure that I had a good appreciation for why that was so significant in terms of the regulatory actions that we do or don't take as an agency.

But, you know, having had more time now to appreciate that a very disciplined approach to that either makes for maybe greater, you know, public understanding of some of the decisions we make, or why we don't take regulatory actions for certain other items that don't make it through the process.

So I watched the animation and how many, you know, managers wanted to come to the microphone when I think you felt like maybe what was being laid out was

some fundamental lack of appreciation on the Commission's part for the fact that at the end of the day -- I'm sorry to have to admit this -- but regulation is not so much an experimental science as it is a theoretical science. I'm sorry, but I just believe that to be the case, specifically where it's nuclear and, as the previous gentleman said somewhat flippantly I guess, but said, you know, "I'm not going to have a core melt."

And some of this on this table are also keenly aware that in the United States our experimental capability and infrastructure to do actual nuclear experience -- experiments with nuclear materials has actually contracted rather significantly over the last 20 years. So, you know, it's simply where we find ourselves, but I -- what matters to me in making these regulatory decisions is that scrutability.

Can people, whether they're our critics or our supporters, can they look in here and see the basis upon which we supported an analysis that ultimately led to some sort of recommendation for the staff to the Commission, and so that our critics can look at that and say, you know, "I think it was either flawed or inadequate." And so that others can say, "No, I think it had a lot of rigor and was well done."

But I just want to have some sort of disciplined approach. And so the convert I guess that you're slowly creating here to HRA is that for all its limitations and inadequacies -- and I feel you have been very candid about where it's limited -- that we at least are trying to have a tool -- as Commissioner Apostolakis has said, it needs to be useable enough and all of the things that tools fall victim to, sometimes being overcomplicated, but -- and if we don't begin and try to use it, frankly, it is our critics who will help us make it better, because they will be the ones that come to us and say, "This isn't right."

And so getting to maybe an actual question, the ACRS -- in addition to Commissioner Apostolakis saying the ACRS's letter was too long, which I'm not sure I agree or disagree with that -- they used wording in here that I have never -- I don't think I have ever encountered wording this strong, but they said that Chapter 7 notes -- this is on the topic of uncertainty -- "Chapter 7 notes that parametric uncertainty in human error probability should be estimated by assuming a log normal probability distribution and applying guidance from NUREG-1278. This is astonishing." That's what they say.

I don't think I have ever heard them use the

word "astonishing."

(Laughter.)

So, but clearly it gets to the topic that has been explored by a number of members of the Commission, which is having some sort of scrutable, high fidelity, if we can have it, approach to these uncertainties. I know that the staff has not yet responded to this ACRS letter, but do you have any initial defenses that you offer to using a log normal probability distribution? It seems rather a crude instrument to me.

MR. PETERS: No, I'm not going to defend that case. This was a draft. We have a recommendation from the ACRS that we are taking very seriously, and we are looking back into getting our team. This is a collaborative team that we didn't really tackle that aspect and just incorporated a current state of practice over into this IDHEAS methodology.

But given that ACRS has a strong recommendation, obviously a strong recommendation to make enhancements to that area, we are working with our industry counterparts to come up with a strategy to solve that issue.

COMMISSIONER SVINICKI: Okay. So you are taking that feedback under --

MR. PETERS: Yes, we are.

COMMISSIONER SVINICKI: -- strong advisement. Okay. Thank you.

And then, again, I just want to say that it may seem perilous, and probably is in some instances, to have to assign a number to human, you know, responses and conduct. In any circumstance I have tried to argue for approaching that in a very balanced way.

And I agree with a number of my colleagues who said, you know, the worst kind of ignorance is sometimes overconfidence that you are able to put something to too many decimal points. But that being said, in order to make regulatory decisions, it has been my experience -- I'm in my seventh year now on this Commission, but, you know, even if it's just a tech analysis that assigns high, medium, and low, you know, what? That's kind of a number. I mean, it's a very crude number.

What I appreciated about Sunil's response was sensitivity analysis, so I often balance where I don't feel that the staff has presented something or they are not -- they can't assign a high confidence value to something, often that is complemented by sensitivity analysis. And I think that's the right thing to do there.

I don't -- I'm not sure what else to do, but it allows me -- you know, being the decisionmaker, which is a specific burden on this side of the table, it allows me at least to say, "How should I weight these?" If this area is both highly uncertain and highly -- of high impact to an outcome, well, then I weight that one way. But if it's highly uncertain and much less significant, then you've given me the tools or you've given me the information that at least allows me to be as informed as I can be.

So I don't think I have any more questions. Would any of you like to react to anything I've said?

DR. WEERAKKODY: I just want to say thank you, Commissioner, especially when you said you got excited about human reliability analysis. I think --

(Laughter.)

COMMISSIONER SVINICKI: Were you hoping to generate at least one advocate or something like that?

DR. WEERAKKODY: I was joking, but I was serious. I really believe, you know, one strength of this agency -- I have been here for 15 years -- is looking far. And to that extent, the fact that the Office of Research is developing these tools, which even me -- we may look at skeptically today, is going to be very useful to us in years to come. So --

MR. WEBER: I would only add, I mean, to your comment that we use all tools available at our disposal to support the regulatory decisions that we have to make is spot on. And we hire the best people we can, so that when we furnish a recommendation to you it's as well thought through and defensible as we can possibly make it.

COMMISSIONER SVINICKI: Okay. Well, I appreciate that. Keep swinging for the fences, Sunil. That's great.

CHAIRMAN MACFARLANE: Okay. Any further comment from the Commission?

COMMISSIONER APOSTOLAKIS: Well, just a quick comment. I share my colleagues' concerns about simulation, but I think -- and I have always expressed those views, even before I joined the Commission -- but I must say the Halden people are doing simulator exercises that are really very impressive. And they do sensitivity analysis on the simulation.

For example, they may give the operators an accident scenario. Then, they hide some information, and let's see how they operate. Then, they do something else. They try to mislead them. So if you look at the totality of this thing, you really learn a lot. Okay? Given the simulation -- I mean, it's simulation, we can't

avoid that -- I'd like to make a comment on the draft report on IDHEAS.

You submitted the executive summary. With all due respect, that's not an executive summary. I tried to understand what the report says. All it tells me is Chapter 3 does this, Chapter 5 does that. That's not an executive summary.

And another thing that puzzled me was to see 40, 50 pages of tables of contents, and I didn't know what to do with them. I mean, giving me the table of figures, I don't know. I mean, you have figures, good. So this is friendly advice how -- what not to do in the future, please. So thank you.

CHAIRMAN MACFARLANE: Anybody else? Further comments? No?

Okay. Well, thank you very much for the presentations and the lively discussion. Thanks to the previous panel as well. I think we are all better informed about human reliability analysis.

And with that, we will adjourn.

(Whereupon, at 12:24 p.m., the proceedings in the foregoing matter were adjourned.)