Official Transcript of Proceedings

NUCLEAR REGULATORY COMMISSION

Title: Advisory Committee on Reactor Safeguards
Reliability and PRA Subcommittee

Docket Number: (n/a)

Location: Rockville, Maryland

Date: Monday, October 18, 2010

Work Order No.: NRC-500 Pages 1-191

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

DISCLAIMER

UNITED STATES NUCLEAR REGULATORY COMMISSION'S ADVISORY COMMITTEE ON REACTOR SAFEGUARDS

The contents of this transcript of the proceeding of the United States Nuclear Regulatory Commission Advisory Committee on Reactor Safeguards, as reported herein, is a record of the discussions recorded at the meeting.

This transcript has not been reviewed, corrected, and edited, and it may contain inaccuracies.

UNITED STATES OF AMERICA

NUCLEAR REGULATORY COMMISSION

+ + + + +

ADVISORY COMMITTEE ON REACTOR SAFEGUARDS

(ACRS)

+ + + + +

RELIABILITY AND PRA SUBCOMMITTEE

+ + + + +

MONDAY

OCTOBER 18, 2010

+ + + + +

ROCKVILLE, MARYLAND

+ + + + +

The Subcommittee convened at the Nuclear Regulatory Commission, Two White Flint North, Room T2B3, 11545 Rockville Pike, at 1:00 p.m., John W. Stetkar, Chairman, presiding.

SUBCOMMITTEE MEMBERS PRESENT:

JOHN W. STETKAR, Chair

DENNIS C. BLEY

MICHAEL T. RYAN

NRC STAFF PRESENT:

JOHN LAI, Cognizant Staff Engineer

GIRIJA S. SHUKLA, Designated Federal Official

SUSAN E. COOPER, RES/DRA

Y. JAMES CHANG, RES/DRA

ERASMIA LOIS, RES/DRA

CHRISTIANA LUI, RES/DRA

SEAN PETERS, RES/DRA

NATHAN SIU, RES/DRA

ALSO PRESENT:

VINH DANG, PSI, SWI

JOHN FORESTER, SNL

STUART LEWIS, EPRI

ALI MOSLEH, University of Maryland

JOHANNA OXSTRAND, INL

GARETH PARRY, ERIN

APRIL WHALEY, INL

AGENDA

Agenda Item PAGE
Opening Remarks 4
Approach for Addressing SRM-061020 on
HRA Methods 5
Break
Qualitative HRA Analysis 84
Adjourn

P R O C E E D I N G S

12:58 p.m.

CHAIRMAN STETKAR: The meeting will now come to order. This is a meeting of the Reliability and PRA Subcommittee. I'm John Stetkar, Chairman of the Subcommittee meeting. ACRS members in attendance are Mike Ryan and hopefully Dr. Dennis Bley will join us.

Mr. Girija Shulka of the ACRS staff is the designated federal official for this meeting. The purpose of this meeting is to discuss the status of the Human Reliability Analysis Methods Development, as part of our continuing interactions under SRM M061020 in November of 2006. We'll hear presentations from the NRC staff and their contractors.

We've received no written comments or requests for time to make oral statements from members of the public regarding today's meeting, but I understand that a representative from EPRI may want to add some comments at some point.

MR. LEWIS: Yeah. I do have a few brief things to say.

CHAIRMAN STETKAR: Good. So we'll make sure we have time allocated for that. The entire meeting will be open to the public. The Subcommittee will gather information, analyze relevant issues and facts and formulate proposed positions and actions as appropriate for deliberation by the full committee.

The rules for participation in today's meeting have been announced as part of the notice of this meeting previously published in the *Federal Register*. A transcript of the meeting is being kept and will be made available as stated in the *Federal Register* notice.

Therefore, we request that participants in this meeting use the microphones located through out the meeting when addressing the Subcommittee. Participants should first identify themselves and speak with sufficient clarity and volume, so that they may be readily heard.

We will now proceed with the meeting. Christina, do you want to say anything?

MS. LUI: Okay. Chris Lui, Director of the Division of Risk Analysis, Office of Nuclear Regulatory Research. As John has indicated at the beginning, that the subject today is to address a staff requirement memorandum issued in November 2006. That particular staff requirement memorandum or notice of SRM essentially, in a funny way, that was issued to ACRS by the SRM. It actually indicated that the SER should work with the staff to address these particular issue.

So even though the SRM was in place in May 2006, because resource considerations from both sides, that really did not start tackling this issue or planning for the current study until late 2008, beginning of 2009.

For example, we were at that particular time that we were focusing on the empirical study, the international empirical study that's coming to a conclusion now, so that we will be able to really start more systematically addressing a lot of the research topics in HRA.

So while the SRM explicitly passing both ACRS and the staff on the HRA model and human performance issues, we do appreciate the committee's interest and input on the work. Clearly today that is a work in progress, given that we are -- we really started working on these issues just about a year ago.

And because the SRM is passing both the ACRS and staff, in consultation with the committee, we plan to have a series of three to four meetings, more in the form of workshops in the next 12 months or so. That way, you will really facilitate the working together and looking for solutions together type of interaction.

The materials that we have prepared for today's meeting are largely based on the interactions that we had back in April this year. In addition to addressing the questions that we have heard and you have asked us to address, we also want a share more complete thought process with you right up front, so it's more transparent to you regarding all the considerations that have gone into the work so far.

As I've indicated, even that it's a work in progress, we value your input and especially those areas where we have identified as challenging topics that we are currently tackling.

At the end of today's meeting, we'll visit what we should be focusing on for the next series of interactions, and we also want to acknowledge that we do have a lot of collaborations with industry.

You see the EPRI representative here, and we also have international participation in the work that we're doing. Clearly, we also are supported by our very capable contractors. So with that, we'd like to turn the table to you, right.

MR. LEWIS: My name is Stuart Lewis. I'm with the Electric Power Research Institute. I'm the program manager for Risk and Safety Management at

EPRI, and that's the program in which all PRA or related activities are undertaken.

I wanted to say just a couple of quick words, since this is really the first opportunity I've had to give sort of an EPRI perspective on what our involvement is in this project. We are, we have been essentially the representatives of the industry in terms of participating with the NRC as a stakeholder in the development of or the adaptation of HRA methods to satisfy the SRM.

I wanted to very quickly give you a little bit of perspective on where we are, from an HRA research perspective at EPRI. So if you could bear with me, I'll so something very quick.

But EPRI did back in the -- especially in the 1980's through the early part of the 1990's, engage in quite a bit of developmental work related to methods for human reliability analysis. The methods that were developed back in the late 80's and early 90's are still in widespread use within the nuclear industry, even though they're a little bit long in the tooth.

These methods included the SHARP 1 framework, which describes what constitutes a human reliability analysis, and identifies the characteristics of a human reliability analysis, including what, at least are the rudimentary aspects of a qualitative analysis.

Then we have two EPRI-developed methods that are still in use for doing further development of human failure events and quantification of those events. Those are the time reliability correlation that you're familiar with, the human cognitive reliability with operator reliability experiments, very conveniently named Method N.

It's compliment or cause-based decision tree method. The cause-cased decision tree method was developed primarily because it was recognized that when time wasn't a predominant factor in determining human reliability, other causes need to be addressed. So that this method was put together. Dr. Parry was a primary author of that method back in the early 1990's.

We really haven't done a lot of work on developing new methods since that time. For about the past ten years, the primary effort related to human reliability analysis is that EPRI has been in the assembly of a tool that we call the "HRA calculator." It's a software tool that's meant to help facilitate performing a human reliability analysis.

Particularly, it was put together because there was a recognition that you can give the same human failure event to six different HRA analysts and get at least six different answers when they analyze the event. So we hope that we can codify the methods to some extent, and eliminate to some degree the variability among analysts addressing essentially the same problem.

I think we've been largely successful in that endeavor at least, that we haven't certainly eliminated variability among analysts or among people that probably view this somewhat differently. But we have certainly reduced it substantially and worked more toward a consistent and repeatable process in performing HRA.

The HRA calculator actually includes a number of different methods for human reliability analysis. It allows the analyst to address both pre-initiator and post-initiator human failure events. Nothing new on the pre-initiator side. It still has elements of the technique for human error rate prediction, and its simplification that was developed for the ASEP program back in the mid- to late-1980's.

The two EPRI methods I just mentioned, the HCR/ORE time reliability correlation and the cause-based decision tree method are the primary tools used for evaluating the cognitive phase of a post-initiator event, and for evaluating execution. The tool does implement again elements of THERP.

So although the methods that are incorporated in the HRA calculator are not very new, the tool itself continues to evolve, continues to be improved in terms of providing additional guidance to users, providing additional capabilities to take more specific aspects into account, and again, continuing to work toward a practical and repeatable process.

As it stands right now, the HRA calculator is in use at every nuclear utility in the United States. That's not to say necessarily that if you went to look at each PRA in the country, its human reliability analysis would necessarily have been done using the HRA calculator.

But all the utilities now have it. It's available by subscription to EPRI. It's not a free product. You have to pay for it, even if you're an EPRI member. But everybody has it, and presumably those who have not yet implemented it are in the process of implementing it.

So it is a -- despite the fact that the methods it incorporates are somewhat long in the tooth, as I said, it is very widely used and widely accepted among the nuclear industry. We think the methods are pretty well understood by the users at the current time.

But we also recognize, looking forward, that these methods are aging, that there hasn't been a lot of review or updating of the methods themselves in a number of years now. In fact, some aspects of what's in the HRA calculator, some of the EPRI methods were set up with the expectation that more work would be done in the future, that more refinement and updating would be done. But in fact, very little of that's been done over the years.

So our motivation for being involved in this process is really to take advantage, as fully as we can, of the work that's being done under the auspices of this SRM response. We think it's feasible to work toward an HRA model that has a stronger psychological underpinning than the models that we have now, and that's certainly been a major effort of this response to date.

We also think that having a more comprehensive understanding of what can bring about a human failure event is worthwhile and something that we should try to incorporate in the way we do business. We also think that it's time, possibly past time that we should be thinking about updating the way we do quantification in HRA.

That's not the say that we think that the methods that we have now are not serviceable. We think that we're getting useful results and insights on the way we do HRA right now. We think that we're doing effective jobs in applying these methods in the context of PRA, as it's used in a variety of applications these days. But we do think that it would be wrong of us not to think about what we could do to make these tools better going forward.

So what we hope that we get out of this, what I guess we're calling now a hybrid model or effort for HRA, from our perspective, we have now a set of tools that we think are very practical for the analysts to implement. I think that's been a very positive development of HRA.

We no longer rely on a fairly small set of HRA gurus to do human reliability analysis. Through the use of the HRA calculator, we've been able to socialize a broader set of PRA engineers in the use of some of the methods, and we think that's a positive thing.

I mean it makes HRA a more integral part of the PRA in general, which has not always been the case over the years. But we don't want to lose that practical nature of performing the HRA. So that's one of the things that we will certainly promote in our interactions on this project.

We also don't expect that this is going to be a method that will require substantial restructuring of PRAs. This again was a, I guess ties back to being a practical aspect of what we're talking about here. There's an awful lot of effort that's been invested in developing very detailed, very extensive PRA models, and we think at this point that the gain that might be achieved by completely restructuring the model to adapt to a new HRA method is fairly limited. So we would hope that what comes about as a result of this project is something that will be fairly immediately useful, without having to start from scratch with our PRA models.

Going back to the objectives that I mentioned before, the HRA calculator itself. We certainly hope that we achieve a set of methods or models that are used as repeatable by different analysts, we achieve a level of consistency in the results and insights that are obtained that are commensurate with what we think we see currently using

HRA tools.

So there are other things I hope, I think we hope to see down the road. But these are probably the major points that we're focusing on, in terms of our interactions on this project at the present time. With that, I thank you for giving me a chance to say a few words about that.

CHAIRMAN STETKAR: Since you're here, I can put you on the spot. What so far is EPRI's role, you know, as an active participant in this project and are you attending methods and providing insights, or do you have tasks that you're working on?

MR. LEWIS: I think our involvement has been a little bit uneven since the project started. Early on, one of our contractors, Jeff Julius, was fairly heavily involved.

Early in the process, and this was before I actually joined EPRI, let alone became involved in this project, much of the effort that was spent on this project was partly the -- are now coming from the Halden research activities, in which it was noted that a stronger approach to qualitative analysis was important for an effective HRA, just worked with the NRC and their contractors to help to refine what that meant, and look going forward.

Starting about a year ago, I became somewhat involved in the project, starting to learn what was going on. I tried to understand how the project moved from what I saw in the SRM to what looked like the development of a fairly new HRA method, which wouldn't have been an obvious outcome to me in looking at the SRM. But I think I understand how things have evolved there, and we're going to hear more from John, as I understand today, about that.

But we have tried to take a more active role in some of the technical elements of the work as well. In particular, bringing Gareth on board to help out has been a significant asset for us. He's been heavily involved in one aspect in particular, in a number of aspects, but one in particular that's been going on for a few months now.

And that is that one of the approaches we decided back in April to take, again moving back to trying to find a practical means to bring all this stuff together, was to use the cause-based decision tree method as a starting point, and we if we couldn't fold much of the research that was being done on the psychological side into the cause-based decision trees, perhaps identify areas where they could be strengthened, modified and possibly where we need new trees, that sort of thing.

Gareth has been working in that area quite heavily. He's also worked with some of the Sandia and INL folks to identify what you'll hear about in terms of proximate causes that might be reflected in those trees, and going back to the factors that could influence performance by the operators. So --

CHAIRMAN STETKAR: Since -- I'll interrupt you for a second. Since you have produced Gareth, is he a contractor to the staff, or are you a contractor to EPRI?

MR. LEWIS: EPRI.

CHAIRMAN STETKAR: Thank you. He's sitting in the corner. It wasn't clear.

(Laughter.)

MR. LEWIS: You know, I think in Gareth's case, that could have worked either way. I think that one thing you can be sure of with Gareth is you're going to hear what he thinks, and it wouldn't change whether he was working for the NRC or EPRI.

CHAIRMAN STETKAR: That's okay. Just curious.

MR. LEWIS: But I think his advice can be used and invaluable as his inputs to the NRC staff have been, through the time we've been working together on this.

CHAIRMAN STETKAR: Thanks.

MEMBER BLEY: Mr. Chairman, I apologize for being late. I was delayed somewhere in the hall. But I should mention, before we go forward, that I am still a contractor to staff on some of the work we do in this area. So in some areas, I have a conflict and I want to offer clarification.

Perhaps in some others I'm not involved in any way, but just as a member, I'd be very careful.

CHAIRMAN STETKAR: Thank you for being careful.

MEMBER BLEY: You can ask him things if you want.

CHAIRMAN STETKAR: Okay. With that, Erasmia, I guess it's yours.

DR. LOIS: Thank you very much, and thanks to the Subcommittee for arranging for us to come and talk to you several times, and work with -- I think we've worked together on the SRM. My presentation will be very high level here today, because Dr. Mosleh and I talk to them mostly, and I talk to them as needed.

So quickly, we'll cover the outline, how we conducted the SRM, what is our approach and aim at the moment, and then the interactions we're going to have with the SRA and what would be the objective of today's meeting, who's involved in the project, challenges, scheduling and what would be anticipated routes.

I'm not going to read the SRA again for you, but I'd like to note, like Christiana did, that we have had several interactions with the ACRS, actually when the SRM was received by the ACRS.

We were invited and also external stakeholders and EPRI came along and were involved in collaborative work, and proposed a plan which we approved and but for several reasons, we didn't start to work until May 2008.

The RES has delayed where we work collaboratively with EPRI, and we're looking forward to ACRS input into the work.

How we interpret the SRM. The SRM used the word "model." We use it more, we interpret it in a more general sense method. So then we believe that we -- SRM is asking to, indicating that a single method is the most desirable, and if more than one should be used, then it should be justified why, which method should be used for which applications, and in that case, we should develop implementation-specific guidance for the specific applications, and desirability for convergence of NRC and industry practices.

So we tried to achieve this, to establish what we call a consensus approach, by developing a quote-unquote "single, high-level method" or factor, which would ensure consistency throughout the analysis process, and be sufficient in general to support application for different domains, and here I have some examples of our shut down, external hazards, Level 2 analysis.

All of these different domains would potentially require adaptation of the overall structure to their specific needs. Also, we hope that we'll gain acceptance from PRA, the HRA and human factors experts, that HRA requires support from both PRA expertise and human performance expertise. This is a very challenging task for us.

We start out with focusing on internal event analysis, and converge with EPRI on that specific area, and then expand to more scenarios and in addition to one of the needs is to be able to develop a screening or scoping analysis for -- to address practicality aspects of some of the various users.

So in April 7, 2010, we met with the Subcommittee, in which we presented a hybrid approach, and ACRS members posed many questions, which are summarized as follows:

Why not use an existing method? Is it a hybrid, why do we propose? How the initial different domains will be handled? Would the approach be suitable for regulatory applications with using existing PRAs? What could be the impact on the existing PRAs? Would it be suitable for new PRAs?

The Subcommittee expressed the need to understand the various facets of the approach, recognizing that there are many facets, and of course how the quantification is going to be handled, and how we will obtain user buy-in. In general, those are the concerns that we've heard, and through informal interactions, we established a plan to have of part of the Subcommittee, so that the members develop, have the opportunity to develop a more in-depth understanding of the technical work, and provide the input, and today is the first of those workshops.

So what we are going to do today, hopefully we'll address what I had before, Items 1 through 5, which show the rationale leading to the proposed approach. Discuss Quantification 6, which relates to the building of technical bases for human reliability using results and inputs from cognitive research and expertise, what we call the mid-layer model.

So that is going to be discussed in detail, if not the first half of the workshop, the second half.

Also, we have, we're going to present another view by, of the current thinking to address quantification, and recognizing that quantification is the area that has several facets. We're going to weigh -- at the moment, we're thinking of a short-term or longer-term approach.

Only the current thinking for short-term perspectives are going to be discussed in the discussion of -- also I would like to point out that the discussion of quantification is in very early stages. We will probably present that aspect.

We would like to obtain feedback and input from the committee members, and plan for the next meeting, topics and Schedule.

So who's involved? As noted, collaborate with EPRI. RES staff directly contribute to the work, and we have NRC-sponsored organizations that are noted here.

The point that I would like to note that this is an interdisciplinary expertise of PRA, HRA, nuclear power plant operations, people who are working in the area and also human factors and cognitive psychology experts. Not all of this are present here today, but actually it's a very big group, about 15 people.

What are our challenges? Well, because challenges are the high-level concepts are appealing. However, when you get to the details, a lot of issues are showing up and need to be resolved. As a result of that, we have an issue to effectively communicate our viewpoints amongst ourselves and we achieve consensus.

As I noted, the project team is comprised of recognized experts in different disciplines, so it's not the easiest work that's being done here. Also, get ACRS input and buy-in, and also facilitating understanding and acceptance by the larger community, both NRC and the industry.

So the schedule, quarterly work meetings with the ACRS. We hope that the technical basis will be ready for public view by September 11th. As the same time, I note here a users guide, and this has to do with what we call before practicality. The technical basis has been intended to be used as the basis for the overall HRA approach.

However, we believe that we can encompass the -- let me rephrase. We don't believe that every HRA analyst needs to be a cognitive psychologist to do the work.

So we hope that we'll develop tools, technical tools, so that the method is being applied in a more -- at a higher level, and yet provide enough basis and be substantial enough so that the analyst has a good understanding of what is being done and how to do it.

Probably we'll be able to also have this methodology ready for evaluation for public review by September. A final report will be ready by 2012.

CHAIRMAN STETKAR: Erasmia, unless you're going to go into more details on this schedule, if you want to leave it to the end, we can also do it then. I'm trying to understand a little bit.

I think I understand the public or the technical basis of that one entailed. Are you proposing to have pilot applications of the methodology in the 2011-2012 time frame, or have you thought that --

DR. LOIS: Actually, we're hoping that

we're doing it, we'll do in a parallel fashion. As part of our overall, of the -- for both approach is being tested or applied, currently with the -- what we call the U.S. nuclear power plant empirical study. But you're right, it may be --

CHAIRMAN STETKAR: Well, I'm talking not just piecemeal parts, in terms of trying to estimate human error probabilities, if you think of that one piecemeal. I'm talking about an integrated framework, in other words, to show how the method would indeed support an actual human reliability analysis, not necessary, you know, 25 actions, but two or three different types of actions within the context of an existing PRA model.

Is there anything -- if you want to discuss this later in the afternoon in terms of going forward, we can do that.

DR. LOIS: Yes. Let's discuss that later, but --

CHAIRMAN STETKAR: But I'd like to a little bit understand whether the schedule has any thought of that type of activity.

DR. LOIS: It does have. We hope that we'll do everything in parallel, but currently I'm --

CHAIRMAN STETKAR: But where in parallel

will that type of --

MS. LUI: Maybe I can give a little more insight into the research. Part of the -- another topic that is going to be coming in front of the Subcommittee next month will be the proposed new Level 3 PRA. So part of the planning is that we're looking at whether this particular method will be ready in time.

If the Commission say that we should go forward and do the pilot on a new Level 3 PRA, is to actually try these method out in the new Level 3 PRA, perhaps in parallel to some of the more established methods.

So we are actually looking at a lot of different possibilities, to allow us to actually use this in a very practical and integrated fashion.

CHAIRMAN STETKAR: Okay. I guess I can understand that, looking forward. I mean there's some uncertainty about the scope and schedule for that activity. I'm looking more backwards, in the sense that Stuart was talking about, that right at the moment, we do have a large number of PRAs, with varying degrees of sophistication and level of detail.

But we do indeed have a large number of PRAs, and a large number of human actions that have been defined with varying degrees of clarity and description, but defined. It certainly would be useful to see how well anything that comes out of this project dovetails with all of those, because in fact, you know, we as an agency are the principal using those make regulatory decisions, and will continue for the existing fleet to use those going forward.

So let's talk more about this again. I want to make sure we have enough time to do all the presentations, because I suspect we'll have quite a bit of discussion on some other topics.

DR. LOIS: So to close my presentation, what we hope is that the methodology to be developed will support several activities such as human evaluation, guidance for staff review of risk confirmed license requests, and for current plans and future plans, and of course would be used for new PRAs and especially Level 3, and industry applications. Okay, so with that, we'll close.

MR. FORESTER: Okay. Well, I'm John Forester, and while she brings that up, I'll just mention that myself and Vinh Dang at the Paul Scherrer Institute will sort of tag team on this presentation.

As you'll note in the title, we are moving towards a hybrid. I think at some of the earlier meetings, we talked about, you know, concerning use of the tool box or piecing different methods together and so forth.

Now some of that may be necessary down the road, but right now our goal is to come up with a hybrid HRA approach. The objective of this is just to inform on approach taken in response to the SRM project.

There's an outline of the presentation. Again, I'll actually just reiterate a slide there, rather than put it up.

MEMBER BLEY: John?

MR. FORESTER: Yes. What's the relationship between the project, NRC and Paul Scherrer Institute in this? How has it been here? It is an interested party or is there some kind of agreement?

MR. FORESTER: As a contractor to Sandia.

MEMBER BLEY: Contractor to Sandia. Yes, I didn't realize that. Thank you.

MR. FORESTER: Okay. Here's just an outline. Again, I'll touch on the interpretation of the SRM and sort of what we're focusing on, which Erasmia has already mentioned that, and talked about our initial process to address the SRM.

What we want to do, you know, part of the SRM directed that we do an evaluation of existing HRA methods. Dr. Dang will give us the results of that evaluation, and then talk about the implication of those evaluations for the SRM project.

Then at the end, I'll give sort of an overview of our current thinking on the hybrid approach, and then Dr. Mosleh and Dr. Parry will end up giving you more detail on some of the key pieces to what we're doing.

Okay. So then this is a slide you've already seen. Our interpretation again, in fact it says in the SRM that probably a single method is going to be the most desirable. So the notion is you have a basic structure, a single method that can be generalized in some way to most applications. And again, there's a desirability for convergence between the NRC and industry, in terms of the use of those very methods. At times, in different kinds of situations, different teams using different methods will come up in different results.

That results in, you know, some degree of conflict or trouble with understanding how to deal with that problem. So if there's a convergence of that and you're confident you have valid methods that's doing what you want them to do, then that would be the ideal, it would seem.

And Erasmia already discussed the last bullet, and I don't think we need to do anything more on that.

Okay. So our initial process for addressing the SRM. Initially, we did do a survey of NRC staff, to see that, you know, what are they doing, using the HRA for in terms of regulatory applications, what kind of issues are they coming up with.

I think we discussed this at one of the earlier meetings. But you know, again sometimes what they find that is they're a method, the industry's using a method and sometimes there's disagreement possibly in terms of event evaluations.

Another point that they wanted to make is they use simple methods that can be practically applied is an important consideration, and that some of the methods also are not being -- some of the existing methods are not generalizing to some of the other conditions, for example, the low power and shut down and so forth.

So again, there is a need for methods that can address these different domains. I think that was --

MEMBER RYAN: I guess I get the largest for one question please.

MR. FORESTER: Yes.

MEMBER RYAN: On the second bullet, desirability for convergence of NRC and industry on HRA methods, is there any kind of analytical approach for that like a certainty analysis or a variability analysis among the models? I guess I'm asking a question. If two models differ, there could be the same answer, but within the variability of the models that overlap for that reason, or they could be different. How are you going to deal with uncertainty, variability and true differences in these sort of models? That's a great question, I guess.

MR. FORESTER: That is a great question. But you know, you'd have to argue that if the methods are valid. I mean certainly, there's going to be some variability from time to time. But if just the methods are generally valid, then they should be getting the same general result.

MEMBER RYAN: I guess my basic question is, and I can understand how you're going to have to deal with that qualitatively, through extra elicitation or other methods, you know, having the right approach at the table. The other is is there an analytical component to that, where you do uncertainty or a variability analysis or an outcomes analysis to help you wrestle with that?

I ask that as a non-expert, but I am -- I'm curious.

MR. FORESTER: Sure, you would try and track how the differences might have emerged, according to the decisions made that led to those differences, to try to get into this thing of why the differences occurred.

MR. LEWIS: But if I could, Ricky, a different interpretation of what -- I'm sorry. Stuart Lewis.

CHAIRMAN STETKAR: No. As long as you have a name tag, you're okay.

MR. LEWIS: Okay, okay.

CHAIRMAN STETKAR: Unless you've changed your name, I mean it's just --.

MR. LEWIS: I'll try and avoid that. I think ultimately what we were talking about in terms of convergence was convergence on a single model, in which case you wouldn't have --- the question of variability among different models goes away. I mean I think that's whole --

MEMBER RYAN: Yes, that's the end point. But I guess I'm curious. Is there, you know, I can understand how you would converge on that by saying well, this is the right model if I raise your hand, and it's kind of an expert elicitation model that you end up with.

The other is you could look at it analytically. If you have three different models and you enter analytic data into any one of those three, are you going to get a different answer? So in that case you could say that all three of the models give you basically the same result. How do you end up with one model is what I'm trying to --

MR. LEWIS: Again, I think it's going to be a process of picking pieces out of different models and some new things.

MEMBER RYAN: But by what measure?

DR. LOIS: We're going to address this issue.

MR. LEWIS: We're going to talk about it. We're going to talk about it.

MEMBER RYAN: Okay, that's coming up. Great, I'll wait.

MR. LEWIS: We'll focus on that criteria.

MEMBER RYAN: Right.

MR. LEWIS: Right.

MEMBER RYAN: Okay. Thank you.

MR. FORESTER: I'll go back one more. I think I just -- we did the survey, and then most of you are aware we had a workshop that involved NRC staff, the laboratories, EPRI staff to sort of obtain their views on what the path forward should be, and there was one thing that came out, what I'll talk about next, is the criteria that we should use to evaluate the HRA methods. There was a general consensus that we'd want to build on the existing methods and experiences that we've had in performing the HRA.

Okay. So here is just a summary of the criteria that we wanted to review the methods against. I'll go quickly through this. I'm sure it's nothing new to most people, you know. We're interested in validity, certainly content validity.

Are we measuring what we know is important, based on what we know about HRA at this point? We have a construct of validity, so we have good underlying models. We're going to look at psychological models as a big part of this. Do we have accepted models that are relevant to the nuclear power plant domain?

So obviously having a method that does a valid analysis is important, and that's one of the criteria we compare against.

Also empirical validity. So how good are the numbers that you get from HRA? What's the empirical basis for those numbers. Obviously, there's not a lot of empirical basis for most of the numbers in human reliability analysis, but some of them certainly have a better basis than others, and they're certainly, you know, not to imply that we don't need to develop better numbers and to use what data is out there.

And one of the key things, of course, is the consistency, the variability and the results, and I'll talk about the empirical study a little bit. But there's plenty of evidence that you find a lot of variability in HRA methods. So we're interested in what we see with the different methods in terms of their variability.

Useability in resources, having a practical tool is really important. So that's something. Again, we wanted to look at the methods with respect, and then look at issues, like what's the scope of the different methods, and also whether they have, you know, an installed base. Do we have a basis from which to evaluate them in terms of their application and their useability?

So this is just sort of a general overview of the criteria. Obviously, we can go with it. We have a lot more detail than that coming out of the meeting.

But so then to perform the HRA method evaluation, we relied on several different inputs. The NUREG good practices document, and the NUREG 1842, where we took, I think, ten different methods or at least nine different methods, and evaluated those methods against the good practices, to see how it, you know, what aspects of HRA they covered.

But in addition to that, at that time we also did sort of theoretical scientific evaluation of the methods that we were looking at. So again, we wanted to examine their strength and weaknesses, and what the underlying basis for the models were, and that's documented in 1842, again for ten of more well-known sorts of methods.

So that's sort of one approach, one type of the inputs we're looking at. We also looked at applications experience and practicality. The NRC uses methods, the industry uses methods. So we have information about experiences from using the different kind of methods, and what kind of problems that people have and so forth.

And then finally we have the results from the international studies. So we have also a proponents-bases HRA method, evaluation approach. Okay, so essentially different elements of what we're going to do, what we used.

For those that may not be familiar, I think we've talked about the international HRA empirical study. But just as a reminder, you know, we have simulator with the Halden Reactor Project, and this was a large international effort, where we had 12 different organizations providing HRA teams, had 13 HRA methods or somewhere in those types of numbers.

Those HRA teams then would take their method and they would look at the scenario. We'd give them information about the scenarios that the crews were going to be performing and what the human failure would be, the conditions and so forth.

Then they had to predict crew performance, and we had 14 crews performing those scenarios in the simulator at Halden. And our goal was to assess the strengths and weaknesses of the HRA methods, and again, we have these predictive analyses, that we can actually see how well the HRA methods performed versus the crew data.

And another big goal is to, you know, try and identify the various, the sources of the variability in the results. So if you see one HRA method that's, you know, several methods come up with this kind of an answer and others come up with a different kind of answer.

We can actually then trace back and see where the differences, where they occurred in terms of the method application. What was it about the methods that led to those differences?

And just while we're here, I'll just mention a few of the major insights from the empirical study. Again, we did see significant variability in results across the different HRA methods. We had -- in one case, we had one method applied by two different teams. We saw variability there too. One thing, one of the main --

MEMBER BLEY: John, just a clarification for the rest of the committee.

MR. FORESTER: Uh-huh.

MEMBER BLEY: You mean both quantitative and qualitative results when you talk about results here?

MR. FORESTER: Yes, I do. Thank you. Yeah, they would --

MEMBER RYAN: Can you give an example of

what you mean by "significant variability"?

MR. FORESTER: Sure.

DR. LOIS: We'll come to it.

MR. FORESTER: It's at the end of this slide.

DR. LOIS: Vinh will talk about it or do you want to talk about it.

MR. FORESTER: Well, if you just wanted to see a picture of the variability, we have an example of the data.

MEMBER RYAN: If it's coming up, that's fine.

MR. FORESTER: Well, it's a backup slide actually.

MEMBER RYAN: If you want to go do it now, that's fine.

DR. LOIS: Yeah, go ahead.

MR. FORESTER: Yes, go ahead. Just jump to the end. I think it made it to the last, one of the last two slides, I believe.

CHAIRMAN STETKAR: John, the final report is out on the Halden --

MR. FORESTER: We have two reports -- we have one NUREG that's done. There's one of the Halden work reports is almost finished. But we still have the overall final report to do.

CHAIRMAN STETKAR: Oh, it's still in progress?

MR. FORESTER: It's still in progress.

DR. LOIS: Both or --

MR. FORESTER: Or we can do one -- this is the loss of feed water. That's one example.

DR. LOIS: Okay.

MR. FORESTER: Now as Dennis pointed out, we also looked at the qualitative analysis that was done by the different teams, and we see obviously that they were done to different levels.

(Simultaneous discussion.)

MR. FORESTER: And they would address different things. So the result of the qualitative analysis that used it as an input into the quantitative analysis varied, and you can also see here this shows the mean human error probabilities, or the human error probabilities predicted by the HRA teams for, what we have four of six different events. Actually, there's only four different events in this study.

The dots in there are the different HRA HEPs. So for Event 1B, as in the loss of feed water scenario, you can see that the box on the far left there says "predicted HEPs." That's the variability and the preliminary probability.

MEMBER RYAN: And that's like an order of magnitude and a half roughly.

MR. FORESTER: Right.

CHAIRMAN STETKAR: Those were all characterized as mean values?

MR. FORESTER: Yes.

MEMBER RYAN: Okay, whatever that means -- .

MR. FORESTER: This actually was a little tighter than what we saw in this steam generator 2 rupture scenario, because there was a little bit more variability there in terms of the prediction that was in --. Do you want to show that one too?

MEMBER RYAN: So these vary a little bit in the range of values, from like an order of magnitude to maybe two it looks like?

MR. FORESTER: Yes.

MEMBER RYAN: And they range in the absolute value of the -- that particular outcome or scenario, whatever you want to call it, by it looks like two orders of magnitude?

MR. FORESTER: Well, but they're different scenarios, right.

MEMBER RYAN: Oh, they're different --

MR. FORESTER: 1B to 2B to 1A to 2A are different.

(Simultaneous discussion.)

MEMBER RYAN: That's interesting. What's the other slide look like? That seemed to be a little different pattern just by breezing through it.

MR. FORESTER: So again, on the bottom are the different HFBs for the scenario, and then the predicted HEPs from the HRA teams. I believe those are ranked along the bottom in terms of difficulty, as determined by the assessment teams, of how difficult the -- how difficult the different HFBs would be given the scenarios.

So you should see some sort of pattern of HEPs going down, and there is that to some extent, I would say.

DR. LOIS: And probably a point of interest here is the fact that those that are on the higher end, which is -- I'm sorry.

MEMBER RYAN: Sorry, you can go ahead.

DR. LOIS: Yes. Well, it's obviously a difficult human action review that we see, tremendous variability among the methods, although analysts were able to assume that it's very difficult to perform.

MEMBER RYAN: That's interesting.

CHAIRMAN STETKAR: And you did say you have one, two groups that used the same method?

MR. FORESTER: We did.

CHAIRMAN STETKAR: And you saw variability?

MR. FORESTER: We did see some variability.

CHAIRMAN STETKAR: Not just in method to method; it's in analyst to analyst using the same method.

MR. FORESTER: At least in that one case, and that's sort of what we wanted to address further in the U.S. studies.

So we studied variability, and one of our main conclusions is that the qualitative analyses that were being done just weren't comprehensive enough, and they often weren't structured well enough for people to use them consistently. So that seemed to be something that we really needed to address.

CHAIRMAN STETKAR: In lay terms, that means people aren't -- practitioners aren't being instructed well enough what to think about?

MR. FORESTER: That's correct, and often what we saw was if you had a really good analysis team, at least your qualitative analysis could be very good, even if the method itself wasn't guiding that. So if you want a methodology that can be used more generally by a broader range of people, you're going to have to have better guidance.

MEMBER RYAN: So there really has to be a lot of process knowledge of the specifics in these reactor cases to get it right? That's what it sounds like.

MR. FORESTER: That's true, and that will vary, again, depending on how much guidance you can give the analysts to use and make sure they ask the right question, the right questions to gather that information.

But I think, you know, having process knowledge what operations in a control room are is always going to be an important part of this. But we'd like to be able to lessen that to some extent. Okay.

Now we're just going to walk through -- there are findings from the evaluation of HRA methods, and Dr. Dang will walk you through that part.

DR. DANG: So the next slides are not going to go method by method, in terms of evaluation. We're going to cover the main conclusions of this evaluation, and before I get into some of those, some of the findings that apply across all the methods.

Basically, all the methods have some significant shortcomings, and they certainly have different strengths. Part of the reason for that is the originally intended scope of those methods. I think the scope of application that we expect of an HRA method has evolved, and some of the methods are stuck back in the days when the scope was small.

The other, of course, relates to trade-offs that were made in the design of the method, I've just given you two examples.

One is simplicity versus the ability to represent a broad range of HFEs and performance conditions. So this is the issue of whether, you know, your method is good enough to measure events full power, or whether you can do local actions, maintenance actions.

Then another example of the trade-off is the repeatability versus the ability of the analyst to consider, to bring in influences that they consider to be important for that particular HFE. The second point across methods is the qualitative analysis, which was mentioned a little bit in terms of one of the insights, if you want, from the empirical study.

This is a shared weakness, and here we mean in terms of the performance issues and the factors that are considered, how broadly they consider these factors, the guidance for this process for the analyst, and then a key issue is the translation, once you find the issues, into the quantification inputs.

Although it's a shared weakness, some methods are definitely better than others in terms of these different aspects, and it's important for us to deal with this, because obviously it impacts the consistency of the estimated HEPs.

When that process is loose, you get different issues being identified, and then whatever method you use, you're not going to get the same results if you're not planning the right issues.

Secondly, it impacts the validity of the HEPs to the extent that this validity can be empirically verified. In general, the data that we have concerning HEP values is relatively weak, so it's less quantitative an evaluation of validity than we might wish. Next slide, please.

CHAIRMAN STETKAR: Let me ask on that aspect, unless you're going to get into this later. I don't want to bog it down, because we covered some of this, I think, on our meeting six months ago.

But I know that EPRI was involved in a number of simulated studies back in the day. It has some catalogue of human error probabilities from those studies that they used to benchmark a particular methodology.

I've heard rumors that other international people, most notably the French, claim that they have a large catalogue. Have you, since you're the European representative -- well, to the extent that Switzerland is part of Europe -- that's a good side joke.

(Laughter.)

CHAIRMAN STETKAR: Have you looked at all, in terms of what other benchmark information might be available on HEPs, you know, because that is a known lack.

DR. DANG: I think the key part of your question is on ATPs, and perhaps I'll leave Stuart to comment on the EPRI data. Also for the -- I mean if we turn to the other large source that you mentioned or large repository that you have mentioned, the French have a lot of simulated data. They are continuing to collect it. It is not so HEP oriented.

CHAIRMAN STETKAR: Is it? Okay.

DR. DANG: So it's more what did we see in these scenarios? What are the crews doing, and they get quite structured in terms of how they describe that. But actual HEPs coming out of simulated experiments, there isn't a whole lot of that. Maybe for them.

CHAIRMAN STETKAR: Are they at all -- are they trying to use that information, though, to structure their own HEP work?

DR. DANG: They use the quantitative information to a very limited extent.

CHAIRMAN STETKAR: Okay.

DR. DANG: They really view is as an input to their qualitative process, to capture the behaviors.

MEMBER BLEY: A clarification issue. And correct me, the French were involved in the Halden study.

They did have a team, and I thought, I don't remember precisely, that after they laid out their qualitative modeling, that they did in fact to some extent rely on what they had seen in their simulators to help in their quantification. But that might not be true. Maybe you know better.

DR. DANG: Well, they also are storing analyses, HRA analyses and their results, and then they're trying to adapt those when they look at it --

MEMBER BLEY: Oh, as data.

DR. DANG: As data, exactly. So it's that sort of data. Now maybe Stuart wants to -- or Gareth.

DR. PARRY: Yes, I can comment on the EPRI so-called HEP data. Those experiments really were opportunistic in the sense that they did, looked a lot of simulator experiments. But there were, I'm not sure that there were any real failures in the observed experiments. What was generated was in fact time reliability curves. So any HEPs that came out of that would have been subject to the validity of those HEP curves.

CHAIRMAN STETKAR: So there too, the purpose of collecting whatever data was collected was a bit different.

(Simultaneous discussion.)

DR. PARRY: It was to verify the concepts for time --

DR. MOSLEH: A quick comment, John. I think you're referring perhaps to the old French data many years back. I'm talking about 20 years or so, where they did collect HEP from simulator exercises.

CHAIRMAN STETKAR: That too, but again, I haven't, you know, I have no personal involvement with what's going in France. So I just hear things, you know, third hand removed at least, which was the source of my comment.

I thought that they said that they still had an ongoing program to collect simulator data, but I don't know what that means. I mean I can say those words, but I don't know what it really means in practice.

The only I brought it up is to explore the extent to which, you know, the collective team has been looking for, you know, HEP benchmark data, if you want to call it that, because that's always an issue when we talk about the validity of the numbers that you're estimating.

Even though everybody might agree that this is the best method possible in the world, if you don't have confidence that it is reasonably predictive, you have problems.

DR. PARRY: In comment, I think the French, it's what's been said. The French are using it mainly provide qualitative insights into the modes of behavior of crews and things like that. Even if they had HEPs, I somehow doubt it would be very easy to get it from the French.

CHAIRMAN STETKAR: That's a different

practice. That's a different issue.

(Laughter.)

DR. PARRY: Yeah.

DR. SIU: John, if I may. Nathan Siu, Office of Research. Vin, I thought you might comment on the WG Risk workshop. Again, did it say what the state of international data collection is for simulators?

DR. DANG: The Working Group on Risk Assessment of the Nuclear Energy Agencies' Committee on the Safety of Nuclear Installations, CSNI, got that through, held a workshop last November on simulated data for HRA purposes.

This is really -- well, one of the motivations for the workshop was to see the extent to which the experience, the very positive experience through the empirical study of collecting simulated data and actually using it for evaluating HRA methods, could be extended and used to support HRA practice and HRA method development.

There was a lot of interest. There remain issues concerning what we really considered data, what we should share sort of at the full basis, privacy issues, transferability issues if you start to share that kind of data. So it's, I think it's an area that there's a lot of interest in, and I think we are going to see some developments in this area. But we're not there yet.

CHAIRMAN STETKAR: Okay.

DR. DANG: Returning back to the slide, we were, to the limited extent that this validity can be empirically verified, I think that in the empirical study we've done a good job of actually managing to do some of the empirical verification.

But there's a limit for the highly reliable actions or the less difficult scenarios. The data doesn't say very much quantitatively there. But you can see there are still some qualitative issues. So empirical validation is something that can be done, and we hope to see more that being done, because it's been very informative concerning that.

DR. LOIS: Before you move to the next, I wanted to bring a piece of information here. I participated in what is called Israel 2010 Conference in Rhodes, Greece, and I presented the outline of our approach here to address the SRM and the EDF representative of human reliability was there.

She noted that MERMOS is pretty much of the same pace and they -- what I was presenting is pretty much what MERMOS is doing. So the point I'm trying to make is that MERMOS is also on a continuing evolution, and they try to address issues related to consistency, variability in the HRA results.

They've pretty much taken advantage of the Halden study. They recognized that they learned a lot and they're continually evolving their thinking towards their simulations.

CHAIRMAN STETKAR: I was going to ask, did you get any sense of sort of if they're proceeding on a parallel basis? Parallel is precisely parallel. Is there some sense of convergence or divergence?

DR. LOIS: They were --

CHAIRMAN STETKAR: Or not enough information.

DR. LOIS: I think we are collaborating and we know what we're doing in terms of exchanging of information. Whether or not we will get to the same point, you know.

CHAIRMAN STETKAR: Okay.

DR. DANG: Returning to findings across methods. So two common practices in HRA that are going to affect this work is the fact that in practice, in many HRAs that are being prepared, you have a mixing and matching of methods.

This goes all the way back to THERP, and it continues today. One split, of course, is decision cognition versus implementation. What we're seeing in THERP, a TRC for the decision part and then tables for the implementation execution part. Practically all the combinations today exist some place.

Another reason for mixing and matching methods are the different contexts. So you have different methods for full power and shut down PSA, just because people feel that on a shut down PSA, with the long time windows, a method is not particularly suitable.

So you're getting already some sense that the practitioner wants to go in a direction of mixing and matching, to meet the needs of the analysis. The second practice is the, what we'll put in quotes, "adjustments of methods relative to the reference guidance, to the documetns."

This is coming from emerging new information since the method was published. One example is people frequently cut off the THERP time reliability curve in terms as you go out towards the long-time windows, because they feel that those low values are no longer realistic, which was not the case when the method was published.

Perhaps what's more key is that there's a great deal of variation in how formalized these adjustments are. In a tool like the HRA calculator, that's a documented set of rules. Use this method and do it like this when you're in this situation, and in other cases, it's really up to analysts' discretion and they actually don't have a document to refer to.

That's the way they've implemented it. They'll document that that's how they're doing it, and that will be the extent of the basis for these adjustments. These practices of mixing and matching and of adjusting the method are going to affect the consistency of the estimated HEPs and their validity.

In particular, when you compare two HRAs using the same method by name, there could be differences and which adjustments they've chosen to apply.

MEMBER BLEY: Vinh?

DR. DANG: Yes.

MEMBER BLEY: You said it will affect the consistency and validity. Did those two go in the same direction, or does making some adjustments help one and hurt the other, or can you decide about that?

DR. DANG: Well, if we're speaking about adjustments to in fact new information, that's going to help validity. If all the analysts decide to do the same adjustment, then we're going to need a better consistency across analysts.

So really we're talking, it depends on whether we're talking about comparing analysts or comparing the results before or after the adjustments. I hope -- I hope that's good enough.

MEMBER BLEY: When I first heard it, it sounded as if I was hearing that any time you make adjustments or mix and match, that you are degrading both consistency and validity, and it seems to me you're surely degrading consistency. Perhaps there are good reasons for those things, in which case one would hope it would increase the validity.

DR. DANG: Right. What I forgot to add as a footnote to these two points is that there are no value judgments being made for these findings.

MEMBER BLEY: That's a good point. Thanks.

(Laughter.)

DR. DANG: That's the way it is. That's what we're dealing with.

MEMBER BLEY: Only observations.

DR. DANG: Okay. So now we move into selected method features, because we're now, you know, we're in the frame of mind of well, what can we use and so on.

When we look at time reliability curves for decision diagnosis, we keep getting this feedback. They're easy to use. they're actually fairly consistent in terms of, you know, once you determine the time window you get similar results.

But they are difficult. What is challenging about the use of these curves is that when you have other influencing factors, it's very hard to adjust what the time reliability curve outputs, produces as an output.

Then some methods are more prescriptive in terms of the analysis, providing work sheets. Obviously, the positive point here is that it supports repeatability, traceability, and it can support validity. However, you need to develop different sets of these rules for each context, and to really consider each context and make all these adjustments in a comprehensive way, in order to have these rules sort of remain compatible and consistent as you move from one application to another.

On the right hand side, it's actually sort of contrasting. We have binary quantification inputs in some of the methods, where you're asking yes-no questions or less than greater than questions.

Clearly, binary quantification inputs are going to support consistency.

You don't have too many choices. It's one or the other, and you don't have to find that gray area in the middle and define it carefully. the negative side is that depending on what you choose on that yes-no, you get sort of a cliff edge effect, and analysts may not like that too much, being forced to really choose. "I don't really fit the yes condition; can I maybe go to the no condition, where I'm going to get an HEP more to my liking?"

In contrast to this type of input are the multiple level, whether it's low-medium-high or full-scale from zero to ten. Here, you can define the input in a more nuanced way, to be able to differentiate. But if you're going to get consistency from such a process, then you need scaling guidance, anchors for the values, so that two analysts will more or less pick the same level on that rating scale, and it should support -- such guidance would support traceability, what is the basis for coming up with a 5 on this judgment, and I already mentioned the task.

MEMBER RYAN: Those are the quantitative markers?

DR. DANG: Yes. You put -- next to the

scale you put what might fit that kind of rating and --

MEMBER RYAN: Yes, thank you.

DR. DANG: Okay. Next slide please. The last one concerns we call it different things, a narrative-based failure mechanisms, and maybe I should what we mean there. Most methods implicitly or in the documentation give you some kind of reference to what failure mechanism is driving a certain table value or a certain analysis.

What we mean by a narrative-based failure mechanism is that some of the methods, and if we look at ATHEANA and MEMOS in particular, these are methods that do this especially, they will explain the failure of a human failure event, in terms of a very specific way in which that's going to develop.

For example, this HFE would fail or could fail because typically, this task is going to be delegated to an auxiliary reactor operator, and he may not be aware of the significance of this information, of some information that he obtains during a task, and communicate it back to the decision-maker.

So that kind of very specific mechanism requires, of course, knowing how the task is performed in quite a level of detail, and the practices of the crews and so on.

MEMBER RYAN: So you're really identifying a weakness in the junior reactor operator's training credentials or something that causes him to misjudge something or make a mistake or --

DR. DANG: I think the point is that --

MEMBER RYAN: How do you translate a junior operator versus a senior into an error?

DR. DANG: Right. More than the focus on the qualifications, it's just the mechanism that once you've delegated, you've introduced the need for an extra bit of communication, that you know, if someone else were doing the task, you would not have this issue.

MEMBER RYAN: Or it could be that a senior operator has to make sure that person's qualified before he delegates it to him.

DR. DANG: Well, we're speaking here in terms of a control room crew, and the tasks that have been delegated are clearly tasks for which the crew is qualified.

MEMBER RYAN: All right. So basically you're saying this person may be relatively new, and even though the training is covered, that maybe it's not locked in?

DR. DANG: Or it may just be tangential information to that task. We're not speaking about just reporting back on that task and the outcome of that task --

MEMBER RYAN: I understand. Thank you.

DR. DANG: So these narrative-based failure mechanisms, they support a very broad set of influences, plan-specific mechanisms, scenario-specific error mechanisms. They require expertise in how these operations are carried out at that specific plant.

This broad set -- this possibility to model so many mechanisms make it more difficult to structure the analyses, difficult to review them in terms of well, why didn't they consider this other scenario, or did they consider this scenario, because there is all these possibilities that are being considered.

And of course, coming up with these mechanisms for an HFE requires a certain level of effort and expertise, an elicitation that is probably beyond what some of the methods, that have the features from the previous slide, would require. Sol there's a lot of effort involved here.

CHAIRMAN STETKAR: Can I ask a question?

DR. DANG: Yes.

CHAIRMAN STETKAR: Something that you mentioned, even peripherally on the third bullet there, where you said not -- analysts might have difficulty understanding whether, at least what I thought I heard, whether there are different scenarios that might be combined into what they're evaluating. Am I understanding that correctly or not?

DR. DANG: No.

CHAIRMAN STETKAR: Okay.

DR. DANG: I mean for example, I think in these methods, you can, of course, you don't have to decide on just one way that all crews are going to carry out the HFE, the task related in the HFE. So you can actually come up with several mechanisms --

CHAIRMAN STETKAR: Okay. So this is really in the context of a well-defined scenario?

DR. DANG: Right. It is within the context of a well-defined scenario, where maybe what I was trying to say that when you see such an analysis, you can certainly judge the possibility of the scenario that's being presented to you or these several scenarios, and say "okay, that sounds reasonable" and so on.

But well, you know, there are so many

other scenarios and maybe across HFEs you see a similar scenario. Well, why didn't you use that scenario here or that kind of mechanism in this HFE. That makes it hard to review.

CHAIRMAN STETKAR: Just so I understand, when you say "scenario," do you mean human response to a specific set of stimuli, or do you mean --

DR. DANG: Yes, yes.

CHAIRMAN STETKAR: --plant behavioral scenario?

DR. DANG: No, I was referring --

CHAIRMAN STETKAR: Okay. Okay, thanks.

DR. DANG: They become --

CHAIRMAN STETKAR: That's where I was getting confused. Okay, thanks.

DR. DANG: So coming out of these method evaluations, both looking at the -- go on to the next slide.

Looking at both the general findings and then what we've learned about some of these features, what we like about them or what maybe gives us some concern, are the following, which become implications for what we want to do in the SRM or what we're doing in the SRM.

One is that the guidance for the quality of analysis needs to be extended and formalized, and that's to make sure that the performance issues, influences, are being addressed comprehensively, and to make sure that this process can be more consistent, meaning that different analysts will cover the same sorts of issues, describe them in the same way, and then provide the quantification of inputs that are similar, when they're looking at the same thing.

The second is the need to support, at least for some HFEs, failure mechanisms in some detail these narratives that I've mentioned.

The third element was rating guidance to support the repeatability and the traceability of the results, and this is now -- once you have formalized your qualitative analysis process, in a way that people are identifying the same issues, we want to make sure that at this interface between the qualitative analysis and the quantitative analysis, meaning the part where you start to use the algorithm to produce numbers, that you're getting consistent inputs.

That's where we believe the rating guidance of the questions, that you need to answer in order to make a particular rating or an answer to a decision, is consistent, and finally to improve the technical basis by linking the HRA to accepted models of human performance.

So obviously HRA methods have considered experience and judgment. To a great extent, the models of human performance, for whatever reason, have been maybe not as emphasized, and in this effort, we were trying to balance these two aspects. That would be important for the acceptance of the HR methods.

MEMBER BLEY: Vinh, from the standpoint of planning and where you go from here, and I guess everybody, one thing just kind of clicked in what you were saying.

We're making an implicit assumption, at least halfway I think it's obvious it's clear, that if we have a good, thorough qualitative description of what's going on in all its aspects, one, we've gotten rid of a major source of disagreements between analyses because we're looking at the same thing.

But two, there's an implicit assumption that we'll be more consistent in our quantitative results, and we haven't actually tried that, as far as I know, providing some kind of detailed descriptor of the situation, and then trying these different PRA quantification techniques on them to see what comes out. Has there been discussion of that?

DR. LOIS: Can I jump in here? The assumption is that the qualitative analysis, the quantitative analysis should be able to reflect the qualitative analysis results. So the quantification, we focus on the quantification to address that and that is the way we want to present our first or initial thoughts on how to do it.

MEMBER BLEY: I might not have asked it the right way. But what I was suggesting is essentially another, I'll use "benchmark" this time, because now we would have a consistent qualitative description for people to apply using their different methods.

CHAIRMAN STETKAR: To meet a consistent, perfect description of the scenario?

DR. DANG: Yes.

MEMBER BLEY: Thank you. Polite response in that sense.

(Simultaneous discussion.)

MEMBER BLEY: And people have perceived the business all laid out, where you might go into everything that applies, so that they'd all be working from the same descriptive picture of what's going on and modeling the same thing.

DR. PARRY: I can't imagine that that

would give you exactly -- that the different methods would give you the same answer, because they don't have --

MEMBER BLEY: But would it be closer than it is now, is the question I have. I hope so, but I --

DR. PARRY: You would hope so, but I wouldn't be convinced, because if the method didn't know how to translate that qualitative information into the quantification --

MEMBER BLEY: It might not.

(Simultaneous discussion.)

DR. PARRY: You can't do it. It might not --

MEMBER BLEY: It won't result in any analyst, you know, deciding well I don't have the tools --

DR. PARRY: Yeah, I'll fudge the answers --

MEMBER BLEY: I'll fudge the answers somehow, for factors of 10 to 100 on our quantitative analysis. That's my question.

MEMBER RYAN: But it seems if you ended up in that place, I would think if, you know, Dennis' scenario was well-defined enough and you ended up in that place, when you see a range of responses, that would get you to, you know, at least some insight into what my question was, which was uncertainty and variability.

And I don't mean variability in an error kind of way; I mean variability among analysts and interpretations and so on. So you know, it would be nice if we could say well, I'll take their brains and say okay, this is the analytical part, this is the interpretive part.

We're not going to get to do that, and I think that Dennis' scenario is very well constructed to the point where we would hopefully, and I recognize it's a hope maybe, but to reduce variability among analysts' interpretations of the facts in the scenario. That would be a helpful step.

DR. LOIS: I think we have done halfway that experiment, Dennis, with the loss of feed or with the Halden study, where we analyzed the loss of feed scenarios, after we had about at least two workshops with all the analyst teams.

Through those workshops, the analyst teams developed a very good understanding of how the tools were performing. In addition to we had the opportunity to communicate the HRA results of the various methods. So then the teams had the opportunity to learn from each other how to do a better qualitative analysis and how to interpret true performance of how the ATHEANA group did and MERMOS.

Both of them did a tremendous job in teaching people how to incorporate those aspects. And yet because the results, the quantitative results that I showed, it shows smaller margins but big enough. The important thing is that even for the most difficult human actions, the variability was pretty big.

So it seems that the quantitative tool constrains the analyst to incorporate the understanding into the quantification, and if you take one decision tree, you have, you're dealing with three PSFs, and then the NRC team had recognized actually what would happen.

But the result was not able to, you know, one, when they did the quantification, were not able to incorporate their understanding in the method. So but definitely we could do more of those studies --.

DR. MOSLEH: I think that obviously part of you're referring to, Dennis, is kind of a control scientific experiment, to show where fixing one aspect of the problem leads to improving the other part, has not been done, only maybe at best maybe partially.

But I can see that through our own empirical validation of what you're doing, we could demonstrate that by improving the qualitative analysis and making it more prescriptive, in the sense that here, we could demonstrate, at least on a limited basis, that the quality of the numerical result moving. That's part of what I think could be the type of test that we should run at this point.

DR. PARRY: Well, back to John.

MR. FORESTER: Okay. Well, I just wanted to give you then just a quick overview of the presentation. Ali and Gareth will both kind of go into more details of what we're doing. But again, we do want to capitalize on the strengths of the existing methods, and also use what information we have on the psychology, the data and models.

So again, there is an emphasis here to build on what we have in cognitive psychology, and, there's a particular kind of activity that we've been trying to accomplish which we're calling now, is to build a human response model, which before we called sort of a mid-layer model.

The notion here is that you use the, you know, there's models from psychology. There's also, you know, explanations about how, you know, cognitive mechanisms can lead to failure. So there's information there, there's data there. It comes from different types of models.

We're going to use that information to help us identify, you know, what kind of human failures are going to lead to the human failure. So the notion there's a lot of ways that failures can occur that will lead to the HFP.

We want to use the psychological models to help us identify what those are, which we refer to as proximate causes, which is just another way of saying, you know, the easily identifiable kinds of failure, like well, they didn't attend to the data, so they never started the problem, something simple like that.

So you use the models to help us identify what are those causes, but then also use the psychological models. If you know what kind of cognitive failures can occur, then you can try and use that information to help you identify what kind of conditions, what plant conditions and the traditional kind of performance shaping factors we consider, could lead those cognitive failures to occur.

So you know, I want to build this relationship between those sort of three parts of it, and again the overall notion here is that that should help us improve the validity.

If you understand what parts of the context could facilitate these failure mechanisms, which in turn lead to the most likely forms of the failure, that should improve validity, and Ali's going to give us a detailed discussion on that.

We also want to use -- and then given that as a basis, we want to use concepts from ATHEANA, which has a very, you know, comprehensive kind of approach really for identifying the contextual aspects that you should be considering in an HRA. There's a lot of guidance there.

To our sense, in terms of wanting to improve consistency, we think that even that guidance as good as it is, in the sense that it covers an awful lot, it probably needs more formal, a little more formalism or a structure that guides the analyst to focus then on the right set of information.

It's not a simple thing to do, but that's sort of the goal of what we think needs to be done, and possibly there's some aspects of ATHEANA and MERMOS-type analyses that we haven't addressed yet.

So again, we want to use those, that part, those concepts and hopefully improve it, and build on the validity and the consistency. Then the next part of this, what we're currently looking at, and again, we're still a little bit of the investigative stage here, but we want to structure.

You know, the emphasis here is in more formalization and more structure. So one thing we're looking at and another advantage of this, of course, is this is what industry uses. If we could take, you know, the kinds of structures used in the CBDT approach for quantification, or some other kind of approach that provides a structure, you know, the decision tree analysis of working it with the Bayesian belief networks, we're going to take a look at that.

But the notion is your quantification, we want to have a structure that we would provide guidance for the qualitative analysis, more structure for that, and then also more structure for the quantification approach. Actually, there should be not a one-to-one correspondence but a pretty direct correspondence between the guidance for qualitative and what you end up quantifying with setting decision trees, for example.

So those things should work, should be together. So again, the emphasis here is on more formalism. It should make it easier to apply, possibly would require less expertise, because again, the type of an ATHEANA analysis without a specific kind of guidance can require a lot of expertise to be able to form.

So again, we think this type of approach, trying to have the structure is going to improve validity, consistency, traceability criteria we see as being important.

Then finally in one of our earlier presentations, we spent some time talking about the crew response tree, this CRT structure, to help us -- there's different views on exactly what they do, but to me at least, they provide a way to represent the different failure paths.

So if you looking at a particular HFE and even ATHEANA puts an emphasis on looking for vulnerabilities, where they might take, you know, what might lead the crews to take a different path through the procedures, well that's one of the goals of the CRT, is just trying to help identify those paths, in again a more formal way, and also that would help contribute to the context for the HFEs.

So I think that's going to be an important piece to what we have. I do think there are still some issues to work out, to see exactly how these work. But I think that's going to be an important piece. But that's only one piece to the whole approach, and again, the goal there will be to improve validity and consistency.

MEMBER BLEY: Can I ask one question about -- I don't see it on the slide, but you were saying you might not need the level of expertise. I mean currently, there's expertise in cognitive psychology. I mean other human aspects of PRA knowledge, operations knowledge, procedures knowledge. Which of those do you see we might be able to cut back on for the guy that's trying to use this method?

MR. FORESTER: Well certainly we don't expect people to be cognitive psychologists. So that's one part of it.

MEMBER BLEY: But we'll provide some help for them on that issue.

MR. FORESTER: Yes. Essentially, we want to provide the information from that, from the psychological models to the model that we use for searching for conditions and quantifying. So that part of it's not required.

I think also, you know, if you know specifically what you're looking for, you may not need as much plant expertise possibly. I think that's always going to be a part of it, and as you -- the more you know about the operations, that that will help HRA.

But again, if you have the specific set of conditions, people are guided --

MEMBER BLEY: Where would they come from, if they don't come from --

MR. FORESTER: Because we'll give them to them.

MEMBER BLEY: The method will tell them the conditions under which --

(Simultaneous discussion.)

MR. FORESTER: No. It will tell the dimensions they need to look for.

DR. PARRY: Yes. I think what John said, the methods should tell them what questions to ask of the right people to get the information that they need to quantify.

MEMBER BLEY: Okay, okay. That's helpful. Thanks.

DR. DANG: And I think there's another type of expertise of which we hope to have less need for, which is if I know that this is the issue, which PSF can I kind of bend to put that into a quantification? By having a better mapping between the quality of analysis process and the quantification, we hope that you don't have to go look for that factor so widely.

MEMBER BLEY: For the human side of this, we're going to provide lots of helps, and somehow on the systems operation side, we'll provide pointers, but you'll still need somebody at the plant or somewhere to help you out.

DR. PARRY: Yes. You still need to know how the plant's behaving and how the parameters are changing.

CHAIRMAN STETKAR: Let me ask a sort of fundamental question that I think I've asked a couple of other times, and I'll ask it again, since we're sort of touching on this right now. One of the difficulties that I've seen is that we take an action, let's say manually start feed water.

I can give that action a name, and you can say "Well gee, under what conditions?" Okay, well under the conditions that all emergency automatic systems have failed. So an operator is required to manually start feed water.

The fundamental question, though, when I look at the billions upon billions, in some constructs, sequences that are created in a typical

PRA model, do I need one human error probability for that that suffices for all? Do I need three? Do I need 50? Do I need 200?

In other words, what elements of the actual scenario development? You didn't tell me that I was supposed to do this under a medium loca, for example, where the scenario response is much different compared to a plain vanilla reactor trip, or that it's in response to the biggest earthquake that I've ever seen in my life, or maybe not the biggest, but pretty doggone large, or a fire.

Where in terms -- so what I'm asking about, is we're talking about this hybrid methodology in terms of trying to provide consistency and validity in terms of quantifying those HEPs that fill int hat blank, in terms of how likely is the operator to fail.

Where, if at all, does this methodology address defining are there one blank to fill in or three or 300 within a PRA model? Some of the early EPRI stuff tried to do that, but not very well.

DR. PARRY: I'll give you my opinion right now. We can't answer it right now, but I think the way it could develop is if once we've developed what I would call a quantification model, hopefully that will give us guidance on how to decompose the HFEs into the right subsets of conditions that you need to address a particular HFE in the different circumstances. I think it should come out of how you develop the PIFs.

CHAIRMAN STETKAR: It might. The right kind of questions to ask. The only reason, the only thing I'll throw out is I'm not a particular proponent of any one method.

The thing that I've always liked about the time reliability correlations, and the only thing, is that it forces, you know, present company excepted, is that it forces the people building the PRA model, the people who are drawing event trees or fault trees or whatever they do to build the PRA model, and the people doing the human reliability analysis, whatever they're doing, and the people doing thermohydraulic analyses, to talk to one another.

Because there are a specific set of questions that need to be answered, like how much time is available. The only way to answer that question is somebody running some God-awful thermohydraulics code for a particular scenario has to indeed run that code, and somebody building the model says "Well hey, over here the time might be different because we've got a medium blow going on at the same time that we failed all feed water."

So that's the benefit, you now, of that particular construct, is that it at least gets different people talking to one another. Now we're trying to say the same thing, but by structuring the most appropriate questions, there might be a feedback.

DR. PARRY: And there would have to be questions about the relative timing of events included in the model. Otherwise, it's not going to represent the human behavior. So probably we are saying the same thing.

DR. MOSLEH: John, on that point, what you're raising is based on the mental and one of the most difficult questions to answer in modeling a representation, and much of the effort so far, I mean a good portion of the effort so far has gone into trying to draw that balance between the various points of view and disciplines that need to enter the picture, in order to identify the right level of, in my opinion, my language, the right level of decomposition, at the level where you capture enough, but it's manageable and practical of the types of the scenarios, the plant scenario and the crew, a response to this scenario.

Why we don't have an answer to offer

today, it's something that we hope that the next time we get together, we'll have some concrete ideas how to kind of basically organize an analysis.

CHAIRMAN STETKAR: That's good. I'm just trying to keep, infuse that sensitivity, that it can't -- whatever you're involved in can't operate in a decoupled environment, because we're talking about consistency in the overall HRA. It's the integrated stuff in the PRA that we care about, that we don't lose track of that necessarily.

DR. MOSLEH: Well, John mentioned that there are a number of pieces to this kind of proposal, and you're going to hear today, or you have heard and you're going to hear more of some of the pieces.

But what the topic, the area, the outline has been the subject of many discussions, many exercises and ideas and concepts that have been in part being tested. But we're not the stage where we have a lot of data.

CHAIRMAN STETKAR: Yes. I think hopefully, I mean if what I'm hearing, and I guess what we'll hear this afternoon, if it's true, in principle, if you have the right construct of questions, without putting it any other way, to ask, they might prod, you know, those PRA modelers, if I can call them that, to go out and search for that right information and therefore define the fact that under, you know, this billions and billions of sequences we need 17 different, you know, 17 different enough constructs, if you will, that require an answer. Okay, thanks. Sorry.

MR. FORESTER: I think that I'm done. We've talked about the next slide pretty much. So the last thing, I think, is we talked about given the size and the practicality of what we're doing, trying to be consistent with the industry, we do want to structure this applicable cost domain.

So again, I don't think, for a particular kind of domain, there may be some special characteristics that you'll have to try and capture. But hopefully the basic structure will work across, will work across domains. So okay.

Then the last slide then the next presentations will be Ali's presentation on the human performance, and again that's what we're formally calling the mid-layer model, and then Gareth will talk about use of the decision tree, to try and see if we can use that kind of structure to improve the approach.

CHAIRMAN STETKAR: Ali, before you start, be aware of the time. Is there a -- I don't know how long you planned here. Is there a break time somewhere --

MEMBER RYAN: Three o'clock.

(Off mic comment.)

CHAIRMAN STETKAR: I'm aware of the clock time. I'm worried about --

DR. MOSLEH: A natural point to break this, or try to fit this in the next 20 minutes? Which one?

CHAIRMAN STETKAR: Huh?

DR. MOSLEH: Which one is your question?

CHAIRMAN STETKAR: Well yes. We have a break scheduled at three, and I just wanted you to be aware of that. Is there a natural break point, or should we break early and then allow you to continue?

DR. MOSLEH: It depends very much on how many questions and kind of what kind of discussion we have. I think that we can do it.

CHAIRMAN STETKAR: Okay. You're not starting where you were last time. You're starting beyond that, right? Are you reiterating what you talked about the last time you guys were here? That looks like you're not. It looks like you're starting where you left off and continuing.

DR. MOSLEH: Yes. You mean the kind of depth we didn't show last time, that we've kind of worked since then, in terms of human performance tests.

DR. LOIS: But before, before, I think we had structured the discussion so that we had your input on this part of the discussion. So if there is no more, probably we should break now, so that we leave that possibility.

CHAIRMAN STETKAR: Let's do that. Let's take our break now. I thumbed through it, and there didn't seem to be a clear break point. So let's do that. Let's recess now until three o'clock, and then we can do continuously until, you know, midnight or whatever.

(Laughter.)

(Whereupon, a short recess was taken.)

CHAIRMAN STETKAR: So back in session.

MS. LUI: I need -- well, while resolving the computer issue, we just wanted to stress that for the second half of today's meeting, that we're really going to dive into a fair amount of technical detail.

So really we'd like to, we'd like to solicit the input from the members. I know that you're not bashful about asking questions.

CHAIRMAN STETKAR: We'll be less shy than the morning.

MS. LUI: And at the same time, the idea of today, we want to get into a position to get your input regarding whether we are at a good position to move forward, whether we have presented a sound technical approach, and clearly, if you have any particular suggestions in terms of path forward, we would like to hear those too.

DR. MOSLEH: Okay. Since the chairman hasn't started the session, mostly as you can see on the title of the slide, the title of slide you see a large team of people. Actually, if I list the number of people that have had input to this process, it would be much longer.

But I wasn't able to consult everyone on putting their name here. I know this is the core team that looked at a particular aspect of this effort. John mentioned that there are a number of facets, aspects, areas that play an important role in what we're trying to address.

This particular one, the team that worked on this particular one is listed. I'm going to be talking about not everything that we have done under the title of establishing a technical basis, but highlight the areas where you feel more comfortable to summarize and present to you and also to elicit your feedback.

There are other aspects that if there is a need and the questions arise in the back of the slides I can go to, and also whatever I need, I will definitely call a friend here, the domain experts, among those present, the domain experts.

So the presentation basically is on obvious segments of what are we trying to achieve, that's coming from the titles that are clear on establishing a technical basis for HRA human performance model.

While we do that, I will explain, and share with you at least a causal perspective I think that is simple, quite actually popular in the HRA and similar domains, and talk about that and use that to guide us what we found in the literature, and which way define the scope of the focus of the literature research in the psychological literature that is summarized here.

Last time we were here, we shared with you a -- or at least provided highlights of some of the results of the search. This takes you a little bit beyond that point, and try to synthesize the information into something that is closer to what can be used in the HRA.

Share with you some findings and examples, so that you get a sense of what you're finding and the developing, and a few words on this, on the expected outcomes.

The goal of this particular activity is really trying to just take an aspect of the set of criteria that we established in the beginning of the project, on needed attributes of the credible HRA, and that's a the contemplated view.

The content validity of this particular domain, like other domains, depends on theoretical framework of the models that we use, methods and approaches that you get to invoke, and particularly obviously this domain we won't do that, to reflect the current understanding in human sciences, relevant sciences and the psychological and human factors and the related disciplines.

Given that where you develop or establish a technical basis, a foundation, a theoretical foundation for the content constructs that you use in your method, you want to basically integrate and then capture the other aspects that are not necessarily obtained from the psychological literature, that we may need to actually extend and go beyond the psychology that we have done and that's kind of obvious, and to go from that perspective.

Just referring to a theoretical basis or a list of theories and ideas that people have published on is not sufficient to establish content validity. You want to also show in more explicit terms how, in a consistent way, how these reflect in the methodology that you're developing or using.

Part of what it does is addressing one issue that we have, we see in the current practice of HRA, what you get from the context of a scenario or a situation, or what aspects of human performance you're considering, depends on a large extent to that theme to develop the method, or maybe individual that is applying the method.

What we wanted to go is kind of broaden the basis, and provide a much broader basis, consensus agreement by tabbing kind of the literature to see what the broader community of experts, you know, psychologists in human factors say about that.

So in that way, you have developed a deeper roots in the community of experts, and then hopefully that helps a consensus, that least in terms of the areas that you're covering in the methodology.

Given that, of course, you know, the next step that you need to take is to formalize the process.

Once you capture the essence of what you -- the findings that you have in the literature, and you formalize that in the form of analysis too, our sense is that it reduces the dependency on experts even in applying the method. In other words, you know, basically you capture the essence of the ideas and methods in a formal, structured and visible way.

We talked about this in the last session, where it's not really sufficient to provide reference to set of information. What we want to do is to see how that actually helps an analysis, and it's not that it would actually remove the need for the expertise, but it formalizes the process that the expertise and information is captured in the analysis.

MEMBER RYAN: Just to understand a little bit, if you don't mind, that formalism would do what? Would it increase reliability, decrease uncertainty --

DR. MOSLEH: It increases the -- decreases the uncertainty in the sense of the existing methods all have some sort of common perspective. But there are differences in perspective as to what aspect of human performance you're looking at. So when you look at the method to method variability, an element of it is the different perspective of the developers.

So that's, if you mean by uncertainty the variability in the results of the method --

MEMBER RYAN: Yes.

DR. MOSLEH: That's one aspect of the model to model variability, an aspect of it can be captured by broadening the base of understanding of human performance --

MEMBER RYAN: So that, I mean so I'm translating what you said a little bit, so I can understand it better. But to me that says that you have a field of outcomes that you're learning more about, that may not become smaller by further study. You may have a wide range of results within that outcome. Is that right? Am I reading that right or understanding that right?

DR. MOSLEH: Well, we were going a little bit kind of maybe lower level of or a deeper level of foundation, in the sense that an aspect -- so since there is variability in the different methods, in terms of their foundation, to move toward the consensus model or approach, one needs to go beyond individual perspectives and approaches, and make that a foundation of the methodology.

That way, once you have this so-called consensus method, it meets fewer challenges form the community. And as such, it provides a kind of a content validity. I'll define the term "content validity" in a bit more detail.

MEMBER RYAN: Okay, thank you.

CHAIRMAN STETKAR: Now just give me an idea. Some methods or two extremes or two examples, one method might say the only thing you need to do is look at every step in an emergency procedure and say what's the likelihood of the operator successfully performing this step or not.

Another method might say the steps in the procedures don't make any difference at all. It's simply the amount of time that's available for the operator to process the information. Those are two very, very different constructs and methods.

I think what Ali is talking about is just trying to see what elements of those things or --

(Simultaneous discussion.)

MEMBER RYAN: --put together to say these actions have to occur within this range of times? Could you have a hybrid model, just to pick on John's example?

CHAIRMAN STETKAR: Well, that's I think

the way they're trying to --

DR. MOSLEH: Yes. The idea is really to cover all aspects, to the extent that kind of the domain sciences allow us, and not to miss important aspects that need to be covered. As such, I think, you know, when you get to the point of analyzing the particular scenario, all relevant aspects, as far as the domain sciences allow us, are covered.

CHAIRMAN STETKAR: Sure.

MEMBER RYAN: Thanks. That helps.

DR. MOSLEH: So the focus of this particular effort is the developing a cognitive framework that links the PRA perspective scenario's context, which we'll define shortly, the psychological processes that are -- that take place in operator mind as well as the crew and team activities, and the resulting performance.

In other words, you go from context in a scenario. You go to the context factors, characterization of the environment through the psychological processes to performance. What you want to do is to do this by, supported by psychological literature, and informed by operating experience.

In other words, we are doing this in the context of the nuclear power plant and particular technology, so what do you find from the psychological literature as to development and meaningful data.

But by capturing such information in a structured way, the information from psychological literature, we think that one of the core criteria would be met, at least partially, and that is the content validity. What is it that you're covering and why, and what is it that you're leaving out and why.

Eventually, if some these things turn up to be maybe too much for the typical analysis, some of these things would go purely into technical basis. In other words, the analyst may not see much of what I say today, but it provides a technical basis to give it much more credibility in the eyes of the development sciences.

So I mentioned content validity. It is really a two core ingredients that we need to address. The content and constructs. The content basically is what is it that you're covering in the methodology, what aspects of human performance? And constructs, and how is it covered and what kind of models and abstractions are used or simplifications you're using in order to address the area or domain that you're covering?

And the fact that there is really, to say the obvious in a way, that consider it a kind of a philosophical perspective, that in order to do that, when we talk about capturing information and from different domains and sciences into explaining or describing a phenomena, we're talking about a causal model or a causal perspective.

So when we talk about human performance model, we're talking about the causal perspective of human performance model, without which it would be very difficult to actually translate what you see and hear in many different domains into something that is useable.

The products of this effort would be ultimately an input to the practical guidelines and the methods for the analysts. But more importantly, at least in this talk, in this presentation, is the technical basis that may go into an appendix to an analyst guide, an analyst would never see.

MEMBER BLEY: Ali?

DR. MOSLEH: Yes.

MEMBER BLEY: When Erasmia introduced everything, she talked about a technical basis document in about a year from now, and a users guide in that same time frame. Are those the two things we see on your slide?

DR. MOSLEH: Yes, yes.

MEMBER BLEY: Okay.

DR. MOSLEH: I mentioned a simple causal perspective -- yes.

MEMBER BLEY: I'm sorry. Just not to pick at how you schedule things, but shouldn't the technical basis be in place some time before you actually were ready to write a users guide for the model, a users --

DR. MOSLEH: Yes, absolutely, and that's what e are pushing, although at the same time, we are mindful of what needs to go into kind of the guidance, in order to see what level of practicality and then simplification we need to have.

MEMBER BLEY: Last question on this. Have you begun writing the technical basis, and is some of that nearing a point where it could be shared?

DR. MOSLEH: Bits and pieces of it, reports that different groups have developed. But they're not in a form that we can -- I mean they have not been internally consolidated and reviewed.

MEMBER BLEY: So the development work's still in progress?

DR. MOSLEH: Yes.

DR. LOIS: So it could be part of the

discussion at the end, what it's going to be discussed in the next workshop, who you call that would be able to provide at least parts of documentation of what we --.

DR. MOSLEH: So I mean there are parts that, if you think of them as kind of draft or working documents that can provide, to get feedback more for the view graphs and the slides.

DR. LOIS: Yes.

DR. MOSLEH: Okay. So this slide shows the simple causal perspective, and it's probably quite a familiar concept to everybody. You see the core ingredients being a human failure event of the type that you see in a PRA or a decomposition of those human failure events that you see in a PRA.

What we want to see next, as the next layer of decomposition and the causal delineation, is identification of a set of proximate causes. The proximate causes are effectively, you know, maybe a formal definition, a characterization of the condition that is readily identifiable as leading to HFE, all right.

So you're looking at the surface and say what specifically happened that caused the HFE, and what we are aiming at is developing a set of proximate causes.

Now the next level of explicit causal model is the performance influencing factors, think of from the context factors, can think of from the generalization of the performance shaping factors. I use the word "explicit" because you see a list of those, and then you'll see a list of proximate causes and a PRA will show a list of human failure events.

What is going to be implicit and part of the technical basis is establishing the relation, the link. On what basis you have a set of proximate causes, what literature and psychology or operating experience supports the proximate causes, and what literature and operating experience supports the list of performance influence and factors.

Such links, the indirect ones, of course, will be soft, indirect and often controlled by the state of the art in the related domains, psychology and plant experience.

To give you a sense of what we call PIF, you can see as a generalization of the PSFs to include the plant conditions. Generally, maybe one can really call these things a context, a set of context factors that are grouped here. Like some of them are plant and scenario-related, time and rate of change of plant parameters, the state of the critical safety functions and components, crew mental model and their assessments dynamically as to what is happening in the plant.

Other things such as goals that they follow, constraints that they might have, and then finally the usual list of human-machine interface, ergonomics, training, other things that are lumped together as PSFs.

Now we do have a list of we're working from, and that list was developed as a result of a dissertation work, Ph.D. dissertation work at the University of Maryland by Katrina Groth.

That work has in it a relatively comprehensive literature review of psychology, human factors and nuclear power plant experience, informed by event assessment and event classification from operating experience, and we have also have three international workshops that provided input to that.

So it was a reasonable amount of effort that went into developing the list of performance influencing factors that we're working from. That list, although it's a starting point, is still something that is going through some modification, as the group works on finalizing the list of things that we need to have.

CHAIRMAN STETKAR: We're probably not going to see that list today then; is that correct?

DR. MOSLEH: We can show you that. It's in the back of the slides. We can -- if you are interested in a --

CHAIRMAN STETKAR: Well, probably not in real time, because I want to make sure we have time to finish today. I was more curious about the current size of the list. Does it number in the one's of events, ten's of events, hundreds of, you know, factors? Hundreds of factors?

DR. MOSLEH: The structure is hierarchical, and it has two things in mind. One is having enough detail hierarchically so that you can help, what Gareth mentioned, kind of asking the right sort of questions in a qualitative analysis.

But also you can collapse them to a higher level of abstraction for quantification, because we know that if you have of, I don't know, 50 or 60 context factors or performance shaping factors, quantification will not be very reliable.

Okay. So we have that list we can share with you later, if time allows.

CHAIRMAN STETKAR: Okay.

DR. MOSLEH: Now let me kind of give you a sense of what we have covered, in terms of psychological review, and this slide is less basically the types of things that have been reviewed, a number of journals, some special issues that summarize many of the findings of the human factors community, cognitive psychology and human factors textbooks, and articles. So the usual search, the usual research in an academic sense.

The methodology followed was to look at -- so how do you extract information from this? Again, if you look at it, after many, many thousands of articles, papers and books that have been published, and then you scan, you have to have a sense of what is it that you want to use and rely on.

One criteria is okay, well there are many people who publish papers, because they need to get, you know, degrees and go to school and all that. So you need to kind of see okay, what level of consensus or agreement is there?

You know, what do you see more of than others? And actually the good news is that there are fewer ideas and concepts and methods than the number of papers written. So we can actually capture much of that in a fewer, kind of a representative ideas and fields.

So the degree of community support or consensus in terms of, you know, uniformity and consistency of the ideas was one criteria, and the other one was the critical, of course, is relevance to nuclear power plant. So you have certain things that apply, particularly you know, people study teachers' behaviors or students' behaviors. That has very little relevance.

So that relevance to nuclear power operation and complex system operation was another criteria.

Given with that, you'll see that there are gaps that are identified. Psychology doesn't support the certain types of things that we know for sure apply to nuclear power plant operation, whether you need to fill those gaps. As such, you end up interpreting, extrapolating and extending from current existing theories. You form a base, you form a foundation, and then you extrapolate from there.

That happens to synthesizing some things that seem to be fragmenting basically. You know, you have different ideas that seem to be kind of pointing to different aspects of the same problem. You go through some sort of a synthesis of those, to at least bring deliberate order into otherwise a very, very kind of a difficult domain, search domain.

Lastly, I think is an extremely important and very difficult thing, is to simplify, to take all these things and then reduce to be more manageable. That's why thinking about guidance to analysts is always in our minds, as we go through this thing.

So now I'm going to look at the causal perspective that we had in two, kind of two sectors. One is on the left side of the earlier graph. The other one is the right half. What do we mean by left side? Left side means going from psychological theories and models and methods to proximate causes, identifying proximate causes of failure from what we see in the literature.

Now the literature, apparently if you look at it, very few -- a smaller fraction of the literature really identifies failure of causes. There are many theories that may provide a hint, and some sense of what proximate causes of human failure might be.

But often, you need to just basically rely on extension and extrapolation from areas where the intent was not in the literature to actually identify failures. So the cognitive models and psychological mechanisms that we identify in the literature were grouped and categorized into initially a list of about 130 and so types, and then further reduced to smaller numbers.

Remember that this is now going from psychology to identify proximate causes, and not necessarily direct plant talk yet, okay. So that's one side of the literature.

In that you can see on the next slide a number of lists, just to give you a sense of the types of methods that we have reviewed, these are titles of these methods, the names. Sensation and perception, information forging theory, situation awareness, sense-making on decision-making and group thinking.

You have the recognition prime decision-making, situation awareness and sense making, broad categories of concepts and ideas on the cognitive engineering and aided decision-making. Again, read the list. These are among the literature that we reviewed, so I'm not going to go through the rest of this thing. There's a --

MEMBER BLEY: Was that previously done early on or this got finished some time ago, the literature review?

DR. MOSLEH: This thing, the initial list that you saw last time, or this may be some earlier, that hasn't changed in terms of volume significantly. In terms of assessment and reevaluation it has, and some of it needs to continue.

One area that we don't have adequate coverage, in my perspective at least, is the purely crew aspect, what if you -- we made a few decisions as to what overarching, kind of a human performance model would be, and then from there, then we see what were the gaps. We saw what were the gaps. We tried to identify the gaps.

And there is a difference between the psychology of the real human performance and psychology in terms of group and crew performance. I think what we have is a little bit lighter on the crew as a team than say on the individual psychology. Maybe you can --

MS. WHALEY: I would agree with that. April Whaley, Idaho National Lab. The literature review is --

CHAIRMAN STETKAR: Please identify yourself.

MS. WHALEY: April Whaley.

CHAIRMAN STETKAR: Make sure our recorder gets the names

MS. WHALEY: Idaho National Labs, and let's see. The literature review is, I would say largely complete, but as we've been replying and simplifying and developing the structure, some areas where we have some minor coverage have become a lot noticeable. So we're identifying, like Alex said, we need to go in and include some information regarding coordination among the crew.

We've got adequate coverage for communications, but not so much command and control issues and you know, delegation amongst the crew. So it's less a broad search and literature review and more targeted to the specific areas that we have unidentified that we're now currently missing.

CHAIRMAN STETKAR: Is that because people have not looked at the crew dynamics, or you just haven't looked for the literature yet?

MS. WHALEY: The latter. There are -- there is research out there about crew dynamics, and it was -- at this point, it was more of we're not sure how to fit this in. So we had to look at it as extensively as we have it now when we're looking at it.

CHAIRMAN STETKAR: I'm sorry. It's about the airline industry. It's a somewhat different crew, but it's still -- there's a history of that dynamic.

MS. WHALEY: Yeah. So there is work out there. Yeah.

MEMBER BLEY: Has anybody tried to take the research in crew performance and dynamics into a human reliability analysis --?

MS. WHALEY: With a specific HRA purpose?

MEMBER BLEY: Yes.

MS. WHALEY: Not that I'm aware of. I could be, you know, have missed something. But it's -- there's a lot of --

MEMBER BLEY: Not that people are doing the HRA development have been aware of that in the picture.

MS. WHALEY: Yes. In terms of how to incorporate that into an HRA and how to quantify it, I haven't gone into that yet.

DR. MOSLEH: One thing that the literature has verified is really how to address the issue of a control impact of continuous monitoring, where you know, what we tend to do in HRA is discretize that into kind of the simpler subtasks and activities, and very quickly you get into kind of the dynamic situation that has not been the focus of many of the task analyses that we see.

Although the literature of task analysis quite extensive, but breaking or decomposition into something that is practical is very difficult to see. Let me comment further on the criteria that we use for developing, extracting proximate causes of failure, human failure from the literature.

One is we wanted to, to the extent possible, have distinct, non-overlapping definitions, and that's with the view on kind of analysts, looking at the analysts and the quantification, having sometimes what we call definitionally --. It's kind of a math term, so that's what we want.

The second one is again, looking at the practical, we want those to be observable or at least inferable, so that when you collect data or information, you see that yes, this is a behavior or a cause that I can associate with failure. I can see it and associate it with failure in a practical manner.

And in a way, you can see that we really didn't impose a top-down structure on the search. In other words, we said look, given the general guidelines, that we want this thing to be plant-oriented, you know, to speak to the technology we're interested, and that we need to this thing to be practical and all that, we asked the domain experts to look at the literature and find what they could find. We didn't impose a structure on that.

So it's kind of, you can call it a bottom-up approach in that sense, okay. But that obviously is not sufficient. We have to look at complimentary, a complimentary approach to fill the gaps, and I mentioned a few of these things before. Use of synthesis models, to take -- we take from multiple theories and converging to one actually unifying picture, and another technique is to kind of read the psychological literature and say well, you know, in this, in the context of nuclear power plant operation, really this really means this, to interpret it.

The other thing it did was kind of try to see if you are looking at a nuclear power plant now, and look at the different aspects of the behavioral response, is there a way of bracketing the problem into kind of distinct set of -- using distinct set of dimensions, to ensure and help us basically assure some level of completeness.

And that is, to give you an example, we all know that there's a certain phase that operators collect information, gather information or respond to incoming information. We said okay, you know, the way you would now kind of characterize that environment, the information, is through three distinct dimension of the problem.

One is the source of information, the source. The plant, procedures, what is the source of information. The other one is modality of the information. Is it auditory, is it visual? What kind of form or shape is the information coming? And the response mode of the operator. Is it going to be an active search for information by the operator, or the operators are receiving the information in a passive mode?

So that, by crossing different kind of dimensions, you can then develop a fairly complete space of characterizing the types of things that could lead to the identification of proximate causes.

So this, together with psychological literature review, helps us develop what I call a list of proximate causes from a generic point of view, kind of in more general terms.

We tried to develop this, such that it would apply to a broader range of applications, not just say a power operation, but also hopefully for shut down, ideally for maintenance and other activities. So keep it generic. Keep in generic in that sense, and keep it generic in the sense of really linking and relating to the psychological literature.

Here's an example of the approach, where you see a literature, excerpt from the literature. We can read the words. A particular reference tells us in this particular case, Endsley's 1995 and references are in the back article, that talks about a way that information-gathering could fail, and that's a problem associated with the way you gather the information, sampling.

You know, you see a set of incoming cues or information, and you become selective or intentionally, unintentionally. For example, from an -- in an inconsistent and incomplete way. What you take that and say well, these then point to a mechanism for a cue not being perceived, or a cue not being attended to.

These key words then lead to identification of the proximate causes that relate to that particular mechanism that I mentioned earlier. So mechanisms being an inadequate sampling, and it says well, if it results in a cue, an important cue not being perceived.

MEMBER BLEY: I'm a little curious about what you're showing us, Ali. Is this like an index card in a research project, and you've got hundreds and hundreds of these from little squibs you pull out of the literature, or is this just an example to tell us this is kind of thing you've done? How has this evolved and how many proximate causes do we now have in a catalogue of some sort, and are they all linked back to the literature like you're showing here?

They all link back to either literature or extrapolation from the literature, in the sense that this describes. So if you look at the catalogue of the least proximate causes, there is an explicit reference to where we have, and then if we didn't have any reference to it, either an extrapolation based on the methods that I talked about earlier. So it's not with a catalogue of mechanisms leading to it.

MEMBER BLEY: Are we getting a bigger catalogue of these kind of things than we had in, you know, one document I'm familiar with, in the ATHEANA document, or are they different or are we learning new things or just a better way to organize them or just a better way to anchor them back to literature of the one form or another?

DR. MOSLEH: Certainly the latter is what might be. But whether they -- I don't know if I can actually tell you what, you know, we haven't done a cross-comparison to the ATHEANA list in a --

MEMBER BLEY: No, I wasn't asking you if you've done that. I'm just trying to get a -- I mean you're showing us an example. But how does it fit into all the work you've done? We were trying to -- we were hoping to get a pretty good picture of where you sit in all of this, what this means.

CHAIRMAN STETKAR: Let me before, just to help me also, on this simple slide here, you accumulated three possible situations, cues not perceive, cues not attended to and cues misread, as all contributing to, I'm assuming, the same proximate cause.

DR. MOSLEH: No, no. Examples of proximate causes.

CHAIRMAN STETKAR: Because for example, cues misread to me could be a much different situation.

DR. MOSLEH: These are separate, yes.

CHAIRMAN STETKAR: Okay, good.

MS. WHALEY: Yes. So these, this is -- the literature excerpt is when we conducted the search, our goal was well, let's identify ways in which or mechanisms for failure in human cognition, and this was one of the pierces that was identified by Endsley, saying incorrect information sampling is a way that failure can occur.

So this is a mechanism, and this mechanism can lead to several proximate causes, which is not perceived, not attended to or misread. So the end result of the mechanism is a cognizant failure, and that's explaining this specific slide.

In terms of how this piece fits into the larger literature search that we've done, if you'll recall from the April meetings, we talked about having a master table of over 130 items. This would be just one of those 130 items. What we found is we have a list of mechanisms were, when activated by the right context, could lead to failure.

They tended to kind of cluster together, and those clusters became -- the proximate causes kind of merged out of those categories. So we have less than 30 proximate causes and we're finding it even further. So when you ask about how many proximate causes do you have, I think we're right about 40.

(Simultaneous discussion.)

DR. MOSLEH: Yes, yes, we have a list here we can show you.

DR. LOIS: You want to touch on it, talk about that?

DR. MOSLEH: So we have separated the

proximate causes in this list of 30, condensed from 130 or so excerpts or content of psychological mechanisms from the literature.

MEMBER BLEY: And the arguments for condensing are that they're essentially identical, or that some other argument?

DR. MOSLEH: Yes.

MEMBER RYAN: Let me ask you another question here on this one --

MEMBER BLEY: Can we get this one done first, Mike?

MEMBER RYAN: Oh sure, sorry. I was -- kind oif it's an add-on.

MEMBER BLEY: Oh, okay. Go ahead.

MEMBER RYAN: This says "Incorrect information sampling." How about correct information sampled incorrectly. Is that a different case? There's two different things I'm asking about. The first one is incorrect information sample. "I looked at the wrong gauge." How about if I looked at the right gauge and read it wrong?

(Simultaneous discussion.)

MS. WHALEY: That would be the misread.

MEMBER RYAN: That's a separate one, okay. But that's counted. Okay, that was my question.

DR. MOSLEH: Yes, yes. If you look at the list of -- there's a few more that I can -- so this, for instance, the one before.

MEMBER BLEY: Do we have a hard copy of this?

MS. WHALEY: No. This was in the -- this is the backup slide.

(Simultaneous discussion.)

MEMBER BLEY: Go ahead and tell us about it, because I can't read it.

DR. MOSLEH: It starts with the, on the left, the types of things that relate to failure in the information perception and process a cue or information not sensed or not perceived, not appended to, misperceived, discounted or dismissed, wrong information or cue attended to. That's the first block.

MS. WHALEY: And we've actually refined that further. We've condensed it further to not perceived, not attended to and misperceived. So we have three.

MEMBER BLEY: Your arguments for this condensation process.

MS. WHALEY: Is that we're removing "not sensed" from the list, because from a psychological perspective, there is a distinction between sensation and perception. But for the purposes of this application, one, you won't be able to identify the difference. It's not observable from, you know, a practical perspective.

So from a practical perspective, you can't tell the difference.

MEMBER BLEY: And you're building a documentation of this process?

DR. MOSLEH: Yes, yes, absolutely.

MS. WHALEY: Yes, and for example, then we're also saying "discounted or dismissed" makes more sense to apply when we're talking about understanding, rather than in gathering information. So we're moving discounted/dismissed to a different section.

MEMBER BLEY: So under something like an information processing model, you're clumping these things where you think they most belong?

MS. WHALEY: Yes, and the wrong cues attended to is actually a specific mechanism or a type of not attending to cues. So it's more specific.

DR. MOSLEH: Yes. That is for -- you have asked about whether we are grouping them. It's the grouping that you see here, for instance, a high level grouping.

Okay. So one group on the left, the upper left, the upper left, you see an I and then you have things that relate now to decision-making and situation assessment, and the right-hand side is the action-related proximate causes, and on your other question he gave you the answer.

But I just want to emphasize the fact that every step of this thing is document, in other words, the links or as explicit as we can make it.

MEMBER BLEY: That will be that technical basis.

DR. MOSLEH: That's the technical basis, yes. yes. There's further translation of this generic list to another list, which is even, you know, more plant talk. In other words, if I go to the next slide that you have on your view graphs, it's on 14.

So on the left you see the proximate called "generic proximate causes." Say take, for instance, a complete -- deciding on incorrect alternatives. So on this list, you have 14? On this list you see two rows. One rows says "In assessing the situation, you may develop a set of alternatives, alternative explanations that may be incomplete or inadequate.

In other words, you know, I look at the situation in a plant and say well, is it tube rupture? Is it loca? It is something else, and then maybe in that list, you haven't really identified all the possibilities, alternative explanations and accident types. So incompleteness in that sense is a problem.

And then once you have a set of alternatives, and the correct alternative is in the list, you make a mistake in actually picking the wrong one from that. So in a generic sense, these are mechanisms leading to incorrect decisions or incorrect assessment of the situation.

So make this plant talk, top talk plant, you need to translate that to specific things that can happen and in effect be cause of proximate cause. It could be skipping a step of a procedure or leaving a procedure altogether.

So what I meant by going through another layer of translation from the generic proximate causes that are identifiable in the psychological literature, to failed cognitive processes, then you need to go to another layer and say well, how does this now translate to specific things that could affect the plant and be a cause of human failure? Okay, okay.

The next step is the other side of the

story. So you have a list of proximate causes, and you have a psychological basis or mechanisms that contribute or would be the background or the contributing factor to cause the proximate cause. These psychological mechanisms can cause the proximate cause if they're involved or triggered by the performance influencing factors.

So the next level of the search from the same literature was to see if we can relate the performance influencing factors to the psychological mechanisms, to the proximate causes, going from PIFs to proximate causes.

So in this slide, you see, for instance, on the left-hand side, the cue not appended to. If you look at the theory or a concept such as situation awareness and the number of mechanisms that are mentioned in that perspective and theory, you can see a certain sort of mechanisms that one can relate to the PIF.

So on the far right, you have the PIFs, three of the PIFs, and then in the middle, you see what the literature or the citation actually refers to, or in this case, narrowing of attention or, for instance, a failure to combine and collapse information or make a final classification to a distinct, recognizable pattern or assessment.

Each of these could point to one or multiple performance influencing factors, the workload, the stress or other things. So the second half of the activity, well kind of happened concurrently, but when you look at the product, you see one leading to proximate causes, and the other one is looking at the psychological mechanism to point to the performance influencing factors.

And so similar level of documentation or literature citation is intended there, although in this case, I would say that the literature is much lighter than the previous case. Here, you have to rely more of our understanding of operating experience and the plant, and things that are commonly known as being performance factors in the HRA literature.

But there is an attempt, and I think an extensive one to the extent possible, is to do the mapping between the performance influencing factors and the proximate causes, based on literature.

So the next slide is a sample of the result of that activity, that you see on the left-hand side a number of the proximate causes, such as cue, for instance, or information not perceived.

Right in the middle in this light blue

boxes, a reference to the literature in terms of the mechanisms involved. Then you see the citation next to -- you might find it difficult to see on the screen, but probably on the slides that are clear.

So in this case, there's three of them that are highlighted. In the next slide, maybe it's easier to read.

MEMBER BLEY: Before you leave this one, the blue boxes are proximate causes. They break down into a subset of some kind of reasons for these occurring.

MS. WHALEY: Yes. I wouldn't say they break down, but these are mechanisms that have been identified as ways at which -- ways in which the cue or information could not be perceived.

MEMBER BLEY: So there's no argument that this is --

MS. WHALEY: I wouldn't say that this is a hierarchical structure.

MEMBER BLEY: You wouldn't? So it's just kind of saying we found this in a bunch of places, so it's important?

MS. WHALEY: Yes. These are mechanisms that have been identified as ways in which --

MEMBER BLEY: And then when you slip over into the context side?

MS. WHALEY: These are the performance influencing factors that we infer or have, you know, been able to directly identify from the literature -- that are relevant for this mechanism.

MEMBER BLEY: Let me take you up to the top row. If those things in the pale blue column aren't a substructure, then why do they go to different context factors, or whatever you're calling these things? If that's not a breakdown, then I wouldn't expect to see the first one go over to four factors and the second one go to to two different ones. So it looks like it's a structure. I'm confused.

DR. MOSLEH: It is a structure. I think what she meant was that, you know, I don't know what the word "hierarchy" was interpreted, but it is really a relation showing the relevance of the factors identified, to say the light blue box, and then relating that to the darkened blue.

MEMBER BLEY: Let me try something else. Stay with the top one. "Cue information not perceived." Then there one, two, three, four, five, six, seven things in the next area, "not perceived due to quality, not perceived due to senor overload."

Now so these seem like different reasons for which it's not perceived. So it does seem like it's structured. Now I don't -- I think what you're saying is you're not arguing that's complete, but those are instances that have been observed, and there could be more instances.

DR. MOSLEH: Yes.

MEMBER BLEY: For those instances that have been observed, you'd go further, and you haven't told us yet how you do that, to link to these contextual factors, which I still don't know quite where they came from and how you do that linking.

DR. MOSLEH: Okay. Go ahead.

MEMBER BLEY: But before we -- it is a structure and they are different, so it's not just citations. It's these are instances that you found, either in the literature or including in event records perhaps --

MS. WHALEY: This is from the literature.

MEMBER BLEY: Strictly from the literature.

MS. WHALEY: Strictly the psychological literature.

MEMBER BLEY: So it's not from any of the event work you've been doing. But it could include things from the event work.

MS. WHALEY: Yes, I mean and this is just the first -- I mean this is entirely based off of the psychological literature, and we're in the process of identifying areas where we need to add operating experience to complete or enhance.

DR. MOSLEH: Let me add on your question, on the pink, light pink side, whether that the list is informed by operating expense or not? Is that --

MEMBER BLEY: No. My question was in the blue one, which are the instances for cue information not perceived, that's a list that you pluck from the literature. That could clearly be -- might be expandable from operating experience. It might be complete. We don't know.

DR. MOSLEH: Yes, yes.

MEMBER BLEY: Now the other question I had was to get from those things that are plucked from the literature over to these contextual factors, I have two questions. One is where do the contextual factors come from, for which you're doing this linking, and how did you do the linking? Or how are you doing the linking? I assume you're not done with all of that.

MS. WHALEY: Yes, yes. We're in the process of that. Move to the next slide. We'll show them. This is the same branch that you were just looking at, only bigger picture, less items.

MEMBER BLEY: It's the same stuff.

MS. WHALEY: This is the same stuff. This is essentially the top brands, the cue information not perceived.

MEMBER BLEY: The right hand side fits in one column instead of two. Okay.

MS. WHALEY: Yes, yes. So it's a little bit bigger font. Okay. When you see the references that say, just that they just refer to number one, that means, that's just entirely our inference.

So based off of -- we have several references saying that the cue or information cannot be perceived due to some aspect of the quality or the availability of the information source itself.

So we have three references there, and the reference is in the back of the presentation.

CHAIRMAN STETKAR: Two of those references having the same author, five and nine.

MS. WHALEY: Yes. Based off of -- if all, on the context if all it says is "one," that's our inference, saying okay, well based off of this, we think that this is going to be a process, and the list of context factors is the list of PSFs that was developed in Katrina Groth's dissertation.

MEMBER BLEY: But you said that also includes plant conditions or Ali said that.

MS. WHALEY: Yes.

CHAIRMAN STETKAR: Was that in Katrina's thesis?

DR. MOSLEH: No, no. Generic, it kind of as a category it does. But if you're referring to specific plants and areas, of course not, because it is -- those are vague contexts specific to which scenario you're dealing with.

Let me show -- maybe we can show Dennis the PIF list. It probably would be difficult. We'll go through view graphs.

(Pause.)

DR. MOSLEH: Dennis, here is the list of the more usual PSFs, but organized through Katrina's work and all that. So if you look at it, it says something like "machine design PSFs." So HSI --

MEMBER BLEY: This is a backup.

DR. MOSLEH: So I'm reading it. This is, so --

MEMBER BLEY: Can't even see this one.

DR. MOSLEH: No, I can't see it. I looked at it for a long time when developing this thing.

It's hard for me --

(Off mic comments.)

DR. MOSLEH: So look at -- the thing says a system response or HSI on the right-hand side, is more plant factors than say things on the left-hand side. So this is organized by organization, organizational factors on the left. Then it goes to the team factors. It goes to personal or individual factors, and then situation PSFs and then machines. So let me read a few examples of each.

On the organizational factors, you have training program, you know, plant programs and activities, resources, procedure, quality and availability, the types of things that are under control of the organization.

MEMBER BLEY: These are the things you brought over from Katrina's work.

DR. MOSLEH: Yes, right.

MEMBER BLEY: Okay.

DR. MOSLEH: And this list was what would I think describe as, has looked at the PSFs that you see in HRAs, whereas we had the three workshops that I mentioned and also the HIRA (ph) data activity that Katrina and April spent a few years on, analyzing a number of, I think, 30 or 40 events.

Then also informed by psychological literature, and then separate one that was conducted by Katrina.

MEMBER BLEY: And if I go back -- given you had this list, and then if I go back to the slide we had before, which was -- there's no number on it.

MS. WHALEY: Seventeen, yes.

CHAIRMAN STETKAR: Seventeen.

MEMBER BLEY: Seventeen. The one up there says you just had this list and you had this citation, and you said seems reasonable that this is the one, one of the things that could be an instance of that particular cause.

MS. WHALEY: Yes.

MEMBER BLEY: Okay, and you had four of them, and all of them are 1's, except one of them was a 1 and a 9, whatever a 9 is.

MS. WHALEY: Nine is a reference to a specific article, or it's in -- basically, every time you see a 1, it means that we made an inference, and if there -- if the literature itself called out, like this is going to be an issue, like a number of the references to Mike Endsley's work, it called out specifically experience and training. So if the literature --

MEMBER BLEY: Oh, I see.

MS. WHALEY: If the literature suggested directly that this is going to be a factor, then we very easily, you know, called that out. If there wasn't a suggestion of what kind of factors could lead to it, then we had to make a judgment or an inference.

MEMBER BLEY: Now for Mr. Stetkar's comment that Mr. Endsley's been cited more than once, I think the goal here is to be as complete as you can as you're doing this. So that's not -- it's not that these are orthogonal references; it's just this is a collection of information.

MS. WHALEY: Yes. These are -- the list at the back of the presentation of references, that's the current references for the literature review. The literature review, as we were trying to get the big models that have, you know, extensive consensus among the community.

DR. MOSLEH: According to the criteria. Now there's kind of the two levels of down-selecting. One was from the literature, to see whether there was a kind of better, more general agreement on the things that the psychologists have said, but another down-selection process was to look at whether it's relevant to --.

MEMBER BLEY: Having skimmed your literature, I have to ask a question. You only cite -- you cite of Reasons' things, but one is his WL, the last book that I know of, on accidents and heroic recoveries. He spends about a third of the book on the good side of human action. Did you try to pick up any of that?

MS. WHALEY: Not for this application, because we were looking for mechanisms for failure. But I think that that --

MEMBER BLEY: I hadn't thought about this until I started looking through his book, and I'm saying gee, you know, he might be right. We're missing -- we don't have -- these don't get reported. So what we have are anecdotal pieces. But anyway, go ahead. You were starting to say something.

MS. WHALEY: Well, yes. I certainly wouldn't disagree with you, that we need to look at opportunities for recovery, and that's something that we continue to mention as amongst our group, within the group, that we need to figure out how we're going to include recoveries in this. So I imagine we'll get there.

MEMBER BLEY: Because we run into a lot of comments from people from plants and our committee, from our folks who came out of the plants, about "Gee, you know. All we hear are operators are bad. What about the other side of it?" Back to wherever you were, say if you can find it.

DR. MOSLEH: Also what to add is that there is that -- so we have the list that we have produced, although they have not gone through kind of all the internal cycles of review and consolidation.

But there is a list of PIFs. There is a list of proximate causes that speak psychology. There is a list of proximate causes that speak plant, and there's a mapping between these, supported by or close to a completed map. But it's something that the team needs to really review.

MEMBER BLEY: The only mapping I've heard so far is I mean the citations from the literature or your own judgment under these. Is there some other kind of mapping --

DR. MOSLEH: Okay. Leaving here, we talked about -- I mentioned the tree element model that I showed earlier, the first few view graphs or somewhere in there, it's kind of basically PIF, proximate cause and HFE, is the surface of the causal model.

Detail I talked about synthesis models.

The synthesis model, the necessity of the synthesis model is because you have a vast group of literature that's different than, you know, different ideas and concepts and RES.

It's hard to synthesize those without having some level of consolidation and the grouping of those into something that would represent at least our perspective of what we learned from the literature.

So those synthesis models provide a way of identifying what we -- recognizing what we have found in the literature, also understanding and identifying the gaps, okay. So I mentioned, you know, looking at information perception and processing; talked about looking at three dimensions of the problem.

One of the dimensions was whether the activity was a passive information-gathering or an active information-gathering. That, the literature obviously doesn't give us that model, you know, as a consensus model. There's a hint of information forging, as is a theory pointing to this.

But we're went a little bit beyond, and that gives us a way of characterizing the proximate causes, in terms of whether they relate information-gathering in a passive way, or information-gathering in an active way, whether the operators are just receiving information or they're looking for information.

So those synthesis models are also kind of under the scheme of this high level construct that I showed you.

MEMBER BLEY: As you say that, I'm trying to envision how these connect and work together in an actual model one day. I guess I don't get it yet. I mean everything you saw sounds reasonable and great, but I don't -- I haven't got the idea of how it all really works. If it's at the stage of really working, or if it's --

DR. MOSLEH: Working in feeding an HRA, yes. I think we're beginning to feed the HRA process with this. I will mention that the focus of this thing is identifying the ingredients that need to be integrated into an HRA process or procedure, and we're stopping at that level.

We can take a couple of steps. Now Gareth is going to talk about how some of these things can be picked up by say a particular way of quantifying an HFE. But the integration of all this into kind of a coherent, cohesive model is something that I can share with you some of he elements. I have a few view graphs. But let me mention this one, so that you see that.

You mentioned earlier or you asked whether your, kind of what the overarching method. Kind of you referred to information processing. Yes, we are using information processing, so that's kind of high level construct, and it's divided into --

MEMBER BLEY: You were telling us that you grouped in something like that, but go ahead.

DR. MOSLEH: Yes. But you are using that, and that is -- we are providing some supporting -- an argument of why that is useful, relevant and it has support in psychological literature.

Once you have that, that gives you an organization of proximate causes. In other words, are these proximate causes relevant or related to the information-gathering phase, or is it relevant to situation assessment and decision-making, or is it really something that gives us a sense of how the actions are executed.

So and there's under each of those, there are also other structures. For instance, some of things that is quite -- it's not a dominant theory. But a lot of people refer to an end book, and that is much of the problem-solving in highly structured environments such as control room environment, when you're procedurized. People are trained and all the things that we know about operator, operating in nuclear power plant operating environment, you know, skill people, procedures and all that.

Much of the problem-solving is really either directed pattern-matching, which is directed by procedure, or operator's thinking about, you know, whether they have a solution that has been basically provided to them through training and operating experience.

So pattern matching or similarity matching seem to be kind of the dominant mode of operator response in that phase of their response. Now that has roots in a number of theories in psychology. So when we talk about finding proximate causes for that phase of operator response, mainly a situation of assessment and decision-making, pattern matching is the dominant one that we're using. So that's a submodel in that phase.

All these things are assembled through the information processing as a model. I mentioned earlier the fact that the information processing model, you can look at it from a crew perspective or an individual perspective, and that's something that we need to finalize. In other words, we are modeling a crew, but going from psychological literature supporting individual information processing concept with crew information processing is something that we need to spend a bit more time on.

MR. FORESTER: I think that maybe one aspect, and I think Gareth will probably touch on it, but there's another aspect of this is that the cognitive mechanisms, when considered in conjunction with the different kinds of PIFs, factors that we expect will influence performance, by looking at those two things again, you get some idea of what you need to measure with respect to the PIF.

So PIF might be training, and you can ask about training at sort of a generic level, how good is their training. But you might also, if you have a mechanism like, you know, similarity matching or some sort of confirmation bias or some sort of psychological process, then it tells you what kinds of aspects about training you need to consider to address that particular mechanism. So it goes a little bit further, I think, and that's not an easy thing to do. But that provides us some of what the observables might be.

MEMBER BLEY: Let me throw in two things. One, I don't see any references to the folks who have done the research on crews in your list, and that is, you're just working on that now is what you said. So none of that made it in.

I see Gary Klein here two or three times, I think, and in your backup you have stuff on naturalistic decision-making. How are you using that?

DR. MOSLEH: Well, so let me go to a couple of view graphs you do not have, and then maybe -- can you go to, I think it's slide -- this is two theories. I will highlight, we will show the slide and I think maybe --

DR. LOIS: Which one?

DR. MOSLEH: It's Slide 32, I believe.

MEMBER BLEY: In the backups?

DR. MOSLEH: The backups.

MEMBER BLEY: We just got your backups.

DR. MOSLEH: Okay, great, good. It says example, frameworks from the literature, and one of them is NDM.

MEMBER BLEY: Yes.

DR. MOSLEH: So it kind of highlights the elements of NDM, and I'll let April comment to which way or if that has been used in forming the list. I can say something about the other one, the recognition of prime decision-making, either one.

MEMBER BLEY: Which is part the naturalistic decision-making model, I would assume, if you call that a model.

MS. WHALEY: There are several -- there are a couple of major camps in the decision-making research in the psychological fields, and the one that is relevant to our purposes is the naturalistic decision-making.

The other decision-making work typically does work with college students, and it's choice among alternatives, and it didn't -- I wasn't the one who did this, reviewed this literature specifically.

But naturalistic decision-making focuses on decision-making by experts in real world environments, and that makes much more sense too, for our purposes, than some of the other judgment and decision-making work.

What we ended up doing is reviewing a couple of the naturalistic decision-making models, of which the one that is by large and far the most published and the most documented is the recognition of prime decision-making, which is Gary Klein's work.

The other models haven't had anywhere near as much discussion and research conducted with them. So what we did with this is the same thing we do with the other literature that we looked at.

We identified the stages of the model, because the model has stages and steps, and then tried to either pull explicitly from the source itself, you know, ways in which failures can occur, and we would take a sentence and say okay, well if this doesn't work, we ended up doing a lot of negation. We take a sentence from the article and change the verbal to "not" or "failed to," you know, to identify a way in which it could fail.

That became part of our big list of 130 mechanisms for failure. It would be part of then our literature for the proximate causes and, you know, connecting to the --

MEMBER BLEY: So somehow those things show up in your current list of 130?

MS. WHALEY: Yes, yes. That was all integrated in.

MEMBER BLEY: It's nice to be able to read how you're putting all this together, because I'm not -- but go ahead.

DR. MOSLEH: In this list on this slide, you see that one aspect is, I think, we think it's very relevant. The other aspect is something that one can model things that way, but I'm not sure if that's really how things happen.

One aspect I think is relevant is that the general strategy followed by people in such situations is matching current situation to prototypical kind of situations. So it's a pattern-matching kind of thing.

The second one, I'm not sure if it applies, is options are considered sequentially. Operators probably don't go from one option to another, in their mind at least, you know. They do consider various options concurrently. But not apparently in the case of firefighters, which was this was, you know, this type of experiment was based on.

So you know, one aspect of it is relevant, the other one is not, and we have to -- that's why I think it's sort of a synthesis model in the sense that I described is important, because we need to extract things that are not part of the theory, combine them with another theory, and step a little bit beyond what the psychology literature is, you know.

MEMBER BLEY: It would be interesting to get into this, because what I remember from this stuff is he's essentially arguing that, you know, it's a snap and I'm going this way, and you do it until it doesn't work and then you snap to another way. We've tried to train people not quite to work that way in power plants, so I'm interested --

DR. PARRY: But they still do.

MEMBER BLEY: But they still do, because it's --

DR. PARRY: They still do.

MEMBER BLEY: Because the first thing in your head sticks, and we see that causing troubles often. But it might do a lot of good too. So go ahead.

DR. MOSLEH: Okay. So I'm going to quickly summarize the last two slides, and then -- so where we are, and this is Slide 19 on your handout or page 19. So we have a lot of the psychological literature that is in place with the few exceptions I mentioned, taxonomy of PIFs that we have, and it may go through some minor quantification, but I think it's reasonable.

The mapping of generic PCs to, those are proximate causes to PIFs, is not as complete as the other reports but it's in reasonable shape. I haven't seen the final kind of round yet from April, but I know that she's been working on it. I've seen it. We have seen an example or two.

The mapping of the proximate causes of generic psychology to a plant kind of specific or plant or a nuclear power plant environment proximate causes, we do have the list or something close to it. The mapping has also been done, but the group has not reviewed that.

And then something that I think Gareth will elaborate more on is basically an example or an area where we think we're heading, in terms of establishing guidance for use of the second part only, in an actual HRA. The one element of that is what I have on the next slide, is consideration of logical structure to gather -- the logical structure is really a model or a mapping between the PIFs and the proximate causes in a form --

I'm talking now about, you know, how an analyst would use this thing, taking the PIFs, going through some sort of a set of steps to relate those to the proximate causes, and then from the proximate causes to PIFs.

In doing so, an example is the decision tree and the variations of that, such as Bayesian belief network, will try to address the relation between PPCs and HFEs, and considerations of what I would call kind of a local recovery.

In other words, if you don't perceive any information, there's a chance to recover from that.

Maybe the next cue will tell you go gather some of that information. So there's what I call kind of a local recovery possibilities that might require a different set of logic to be imposed on the model.

There's the interdependencies potentially between the proximate causes, because of the common PIFs that apply to, and the PIFs are also not independent. They have interdependencies that need to be considered, at least, in an explicit way.

And some of these things we hope that will be subject of the next timer we meet to talk about some of these structures, although Gareth will provide a hint of this as to what we are --

MEMBER BLEY: You dropped the phrase "Bayesian networks" in there somewhere.

DR. MOSLEH: Yes.

MEMBER BLEY: I didn't quite get where it fit and where you were thinking it would be helpful.

DR. MOSLEH: Yes. Well, the PIFs are interdependent, and then when they map to the PCs, also there is sometimes a common or set of common factors that point to several PIFs.

When you're facing a situation like that, when you know that the influence path is complex, and there are statistical and causal interdependencies, then the most natural model, the style of modeling I know is influence diagrams. The next step would be the Bayesian influence diagram.

Whether we use the Bayesian belief network to explicitly quantify is something that we haven't really explored.

MEMBER BLEY: Fair enough.

DR. MOSLEH: So the last one is some of the team members are also team members that have current data activities, and that's -- so we're trying to coordinate, at least in terms of exchanging information, in a sense coordinate with the activities such as the one that you're familiar with, and that's similar to data collection as a minimum these two projects.

MEMBER BLEY: For the other members, that's time lines for real events and --.

MR. FORESTER: Gareth.

CHAIRMAN STETKAR: Okay. You're up, sir.

DR. PARRY: Okay, thank you. I'm going to talk about developing an HRA quantification model.

(Laughter; simultaneous discussion.)

MEMBER BLEY: The thing is it's spelled correctly on this.

DR. PARRY: I know, I know. But my email

address is incorrect, so drop the 47. Is that right?

(Simultaneous discussion.)

DR. PARRY: My name's spelled correctly. So Ali's talking about decomposing everything into finer and finer pieces. What I'm going to talk about is the approach we're thinking or an approach we're thinking, of stitching it all back up together again, to come up with an HRA quantification model, okay. So this is the model that we would use.

Go to the next slide. This is -- and what we mean by an HRA quantification model is that the model that we use to generate the HEPs for a well-define human failure event. So that's what I'm talking about, is how we're going to develop that model, and this is one concept for doing it.

The idea is that the model going to use, as the input, the performance influencing factors and we've used that term a lot, that define the overall context for the HFE, okay.

The context is in two parts. We're going to use the plant conditions; it's what's going on in the plant, and also the traditional PSX that typically you don't get directly out of the PRA scenario definition. You have to learn about them from talking to crews and looking at the procedures.

I've got here four like sort of ground rules in this presentation. They're certainly things that I think we would like to have the model look like.

First of all, the whole basis for this project is that we want to have the model to have a sound basis that's consistent with cognitive psychology as a behavioral science discipline. So that's the reason we've embarked on this SRM project.

The second thing is that we'd like the model to be practical and applicable directly to human failure events that are defined for PRA models. The third thing is that, and I think this is important point, is that the model should provide the user with the tools to identify the crucial elements of context that you need for the quantification.

So in a sense, you could almost say that the quantification model developed as it is based on all the work that's been done here, should provide for you the guidance for the qualitative analysis that you need.

The fourth thing is that we want to have a model that will enable both intra- and inter-analyst consistency. Okay. So we want the model to be as objective as possible, if you look, in some sense, given the context.

Okay. The theoretical basis for this quantificational model or quantification model is the -- I believe you would have heard about the nested information decision action model before.

Okay. The concept behind that really is that the HFE is the failure of a process which is not a point in time process, but it's actually a distributed in time process that if, you know, something happens at the plant, if somebody responds to plant changes.

This is an opportunity to recover, as you mentioned Dennis. We have to worry about the recovery factor, and it's the failure of that total package that is the HFE. So we want to basic go on the sets of proximate causes that we're developing as part of this project.

And the justification for the quantification tool is really the set of cognitive mechanisms and their links to the PIFs, and back to the proximate causes.

MEMBER BLEY: Gareth? Maybe I'll wait until you go a little further, Gareth. But eventually I want to -- I'll telegraph it. Eventually, I want to ask you about the IDA and the nested natural, and I've seen Ali show that several times.

The one place I have a little trouble there is if we're using this in the modeling, the "D" is a big piece. I mean the D has got the kind of mental model and situation assessment. It's got the diagnosis and it's got that plan formulation kind of stuff all embedded in it, which is a lot of stuff going on.

So maybe later I'll raise that, but I just wanted to telegraph it. But if we're using IDA and these recycled loops on it.

DR. PARRY: Okay. I'll explain that now.

MEMBER BLEY: Then I'll want to know a little mor about them.

DR. PARRY: I don't think, I mean at least in my mind, that I wouldn't use that explicitly. I'd use that concept, that thought process implicitly in designing the approach for quantification.

MEMBER BLEY: Okay.

DR. PARRY: But it wouldn't be an explicit loop type model, but I think that it's really too complicated. Okay, the next slide.

The approach is that for each of the proximate causes that we've decided, there's two constructive decision trees. Okay, this is -- if you're familiar with the CBDT, the old EPRI project, it's similar in concept to that. The trees are almost certainly not the same. That was after all done very quickly on the basis of some observations from the EPRI experiments.

But the idea is to construct trees, and the headers on the trees will represent the critical PIFs. So that meets ground rule one, if you like. But it helps me with ground rule one, which was to have a sound basis that's consistent with the cog psych and behavioral science discipline, okay.

Then for each of the branches, the aim would be to have a set of questions provided to help an analyst determine, as objectively as possible, the existence or not or PIF. So I think this gets to a question that you asked, Dennis, about how do you make sure that the expertise from cognitive psychology and also from the plant people, get into the quantification.

And that's the way to do it, is by having the analysts ask the right questions of the right people.

A comment on these and what these decision trees represent is that each path through the decision tree essentially represents an explanation of why that human failure occurred, why that proximate cause occurred, in terms of both the cognitive mechanism. Now that, I've got to say the cognitive mechanism won't be explicit on the trees, okay. But it will underlie the structure of the tree. So in terms of that and the PIFs.

This is not dissimilar to the MERMOS approach, where they create what they call little stories. But basically they have a mode of operation of the crew, which could be that they're stuck in this particular mental model, for example.

They have performance influencing factors, but they call them something else, which would enable that to exist. But they also have a very important thing, which is the failure to reconfigure their mental model, which is the recovery factor.

I think the idea would be to have that recovery factor built into these decision trees, because it's really important in many ways. To give you an example of that would be if there was an issue related to failure to detect a crucial piece of information. That would be perhaps the thing that would kick it off, that they failed to see some cue.

But one of the recovery mechanisms could be that there's an alarm that comes in late in the day, that still gives them enough time to recover. So you want to include that also as a recovery mechanism. So that's the type of thing that we would build into the decision trees. Okay next.

MEMBER BLEY: Are you envisioning a standard set of decision trees, or perhaps a set of branches that, depending on the situation, could be reassembled in a tree that fits the particular situation, or have you gotten that far along?

DR. PARRY: Well, I think my vision is, okay, and I'm not sure. We haven't all discussed this yet, so it's not perfectly true, but I think my vision would be that we would have a standard set of trees that would be a, if you like, consensus model of what this quantification of models should look like.

But that the, for a specific HFE, then the answers to your questions would define which path you were on.

MEMBER BLEY: Okay. So you'd be like pruning this thing down to --

DR. PARRY: Yes. It would actually take one path through the tree for a particular HFE. Now if it was -- if that HFE and this gets back to your question, John. If it was an HFE that occurred on the different context, then maybe you'd go down a different branch, okay.

So that the trees would be as complete as we could with the knowledge we have, but that they -- it would be the conditions and the assessment of the PIFs that would drive you to the right end point for the particular --

MEMBER BLEY: Let me dream just a second. This thing could be very big.

DR. PARRY: It could be very big.

MEMBER BLEY: But you could have a set of questions that as people answer it would automatically simplify this thing, to get down to what you're looking for.

DR. PARRY: I think that's something that we've got to be aware of, that it could be very big. Actually, for the few examples that I've tried to construct so far, they're not that big.

MEMBER BLEY: No, okay.

DR. PARRY: Okay. They're bigger that the decision trees in the current EPRI project, but they're not enormous. Now the approach I think that we're going to try to come up with initially to quantify these trees, so far the only trees I've constructed, what I've done is I've ranked the end points according to my perception of what the HEP should be.

Okay, so the place -- well meaning the one with the worst conditions would have the highest HEP, right.

MEMBER BLEY: Right.

DR. PARRY: The one with the most optimal, with the optimal conditions would have the lowest HEP, and then each of the end points on the decision tree would be ranked, okay. That's a ranking we'd have to agree on, obviously, amongst ourselves.

The idea to use expert judgement to come up with the HEP for each path through the tree ultimately, okay. Now that's not going to be easy. But particularly because, as Ali suggested, that the PIFs are not independent, and furthermore, we don't have an empirical basis really for judging the degree of dependence.

MEMBER BLEY: So whatever we come up with is going to be changed, I suspect.

DR. PARRY: It could change.

MEMBER BLEY: It could do, yes.

DR. PARRY: Well, that's true, but I mean I think that's -- that aspect of it should be embedded in the trees, okay. So the idea is that these expert panel that judges these probabilities is going to come up with these numbers. I see Dennis smiling.

MEMBER BLEY: I won't say out loud what is crossing my mind.

(Laughter.)

DR. PARRY: Okay. I mean it's going to depend on which experts you have to do the trees, right? So I think the only way to get consistency is to have a consensus approach to this. So rather than have each analyst that's applying this model come up with his own probabilities, these end points should have probabilities that have been chosen by consensus, by some group of experts, okay.

I mean that's what was done with the original EPRI CBDT effectively. The numbers were just chosen. That, in my mind, is the only way you can have a consensus model. Otherwise, we're back in the situation where everybody who applies it is going to apply his own rules, which are going to be very difficult to unravel, to understand why they got the numbers they got.

Okay. So the proposed approach, at least originally we're going to give it a shot anyway, is to rank the end points by the HEPs, and remember what the end points are is this is the HEP under the set of conditions, that's defined by the path through the tree, okay, by the combination of the PIFs, and then certainly reach consensus on two bounding cases for this tree, and somehow interpolate.

So that's one way of reaching a consensus model, in the absence of a database that we can use to quantify these things. We know we don't have one, okay. So let me just make a couple of comments about what this quantification model results in.

It's effectively, if we were able to do this, if we are able to do this, and I believe we can up to a point, then what it will be, it will be a consensus quantification model. That's consistent with what we've learned about the psych literature.

MEMBER BLEY: Not to pick, but when I look back to find ground rules one and two and four, all I find are ground rules one and three on your slides.

DR. PARRY: That's all?

MEMBER BLEY: That's all I find.

DR. PARRY: Oh. Not on the one I've got. That's on page two.

MEMBER BLEY: Ahh. They started sooner than that.

DR. PARRY: Oh, those are the --

MEMBER BLEY: I'm sorry. I was picking them off later slides.

DR. PARRY: The whole set's there in the --

MEMBER BLEY: Yes, the whole set's there. Pardon me.

DR. PARRY: Okay, so -- okay. So the idea is that, you know, that this model would be a practical model to use, and give guidance to the -- actually it satisfies probably two, three and four, I think. But the -- well, it satisfies all of the ground rules.

But the way you've used this, okay, is that the model will define the minimum amount of qualitative analysis you need to apply -- to estimate the HEPs, okay. Some of that comes from the PRA scenario definition of the HFE. Some of it's going to come from looking at the procedures, talking to trainers, talking to operators, for example.

I think this model could be used to directly for existing PRAs. It could also be used to support an ATHEANA type analysis, where you look for deviation scenarios, because you would just -- you might have for one HFE in a traditional PRA, you might have a number of different ones with different subcontexts, if you like, for an ATHEANA-type analysis.

So you would just, for each of those subcontexts, you'd have a different pattern of the tree, but that it would be weighted by the probability of that subcontext within the PRA context, the error fault context in other words.

I think if we structure these trees well enough, based on all the information we have, the trees themselves can be used to provide guidance on how to choose the subcontext, how much to subdivide the conditions, if you like, to give you clearly different subscenarios.

So I would envisage that the model could be used iteratively with a method like the CRT or ATHEANA, to develop a more detailed HRA/PRA model. It doesn't need to be done that way. You could even do the HFE online, off line I mean, by just looking at the different subcontexts and weight the different HEPs you'd get for the different conditions.

So and the way this model would work is that for any scenario, PRA scenario with an associated context, then the HP of the HFE, given that scenario, would be the sum over the probabilities of the proximate causes given that scenario. Okay. That's the way we're suggesting that this model works.

The second example is for a place where

the context is broken down into the subcontexts, this just becomes a double sum where, I mean it's a straightforward extension. So that's the idea behind the model. I think the other advantage of using this type of model is, and this advantage is not dependent on whether we eventually choose a decision tree type approach or any other type approach.

But we are leaning towards using a causal approach, and any causal approach is probably going to give you a more rationale basis for looking at dependencies between HFEs and the scenario.

MERMOS does this by looking at the different relationships between the different what they call human failure scenarios, that are different ways of failing the HFE, if you like. So they look at the perpetuation of the modes of crew operation into another HFE further down the road, for example.

So that's the general discussion. I did have some backup slides of a very crude initial shot at a decision tree, if you want to see maybe what it would look like, just to whet your appetite of something.

MEMBER BLEY: Why don't you show one? Right now, it's so vague to me, you know. It smells good.

(Simultaneous discussion.)

DR. PARRY: Could you back up? I want to see which there we're looking at. No, no. Back up. Yes, I know. This one, okay. Okay, yes. The particular proximate cause in here is that the crew intentionally dismisses information, okay. So in a small subgroup --

CHAIRMAN STETKAR: Are these in the backups slides we already got?

DR. PARRY: These are different backup slides.

(Simultaneous discussion.)

DR. PARRY: In a subgroup of the project team, which is John, Katrina, Stacey Hendrickson and myself, we set out looking at a particular -- it's a slightly different set of proximate causes from the ones you've seen. I mean these are the issues we've got to work out obviously.

CHAIRMAN STETKAR: You've got to start somewhere.

DR. PARRY: And I think that came out from work that Katrina and Stacey have primarily done, based on the work of April and Johanna. So these were the PIFs that we decided were, that were decided were the appropriate PIFs for this particular proximate cause, given what we could find in the psychological literature, okay.

So one of the things -- well, one of the reasons you might dismiss information is from bias. One of them could be conditioning events, past experience with a specific indicator could tell you that it's not reliable.

I think in a lot of these information processing things, ambiguity of system responses seems to be a key factor, because I think you have to have some ambiguity; otherwise, why would you even tend to make a mistake. So another one of the traditional PIFs obviously is training, which will induce bias, if nothing else.

And then there's knowledge and experience, okay. So these generally are the PIFs that we were working from. So based on that, and I can't even read it now; I'm sure you can't.

MEMBER BLEY: You're right. I can't.

DR. PARRY: Okay. So this was an attempt to put all that into a decisions structure, okay. So the first branch point is is there a mismatch with expectations? Okay. So this brings in the bias aspect of things.

The next question, and don't ask me to

explain why they're ordered this way right now. This was -- I mean obviously it's --

CHAIRMAN STETKAR: It's a complete, logical structure.

DR. PARRY: Yeah, okay. The things that we have are the indications are unreliable. One of the questions we asked is are they confirmatory information indications. This would be a recovery mechanism of sorts.

So the final question is the recovery potential. If there is confirmatory indication, is there a reason why the crew would recognize that that is going to put them on the right path? I mean there is procedural direction to take account of that, for example.

So that's the idea of the structure. Then at the end, you've got these rankings here, and basically the way you've structured the tree is that the bad conditions always go up, and the good conditions go down. So they rank from one to ten, with a negligible stuck on the bottom there, because I think wow, you know. This is a perfect situation. Why would he think of dismissing this information?

And so the way to -- remember what we said is that we want to have objective guidance on how to determine whether you go up or down on these branch points. So this is an example of one of the branch points, and again don't take it too literally yet.

But the idea is this is the branch point that's concerned with confirmatory information, and can the significance of missing information be validated by other information? Okay. This is the significance of the information that they've dismissed.

MEMBER BLEY: I know this is just an example, but from the earlier discussion, it would seem you could have questions that come from the other side, like the Gary Klein stuff. Are you locked into an initial choice here and that's one of the reasons why you might just keep going or something like that, so you could build a larger set of questions that do both sides of the issue, that drive you?

DR. PARRY: Well, yes. I think that locked in one would come earlier on.

MEMBER BLEY: Oh in the way you compose the tree? That's fine.

(Simultaneous discussion.)

DR. PARRY: Right, because this is allowing the possibility of recovery. So that the -- this question of whether it's confirmatory information would -- you'd ask things are there additional indications that would typically be used to confirm the plant status.

MEMBER BLEY: The place I'm nervous here is in real events we've seen. There are cases -- if you just put these kind of questions, and there are things, well then yeah, they're going to --

CHAIRMAN STETKAR: Yes, it is. It's there.

MEMBER BLEY: They're going to recover. But you could also have cases where you've got a mind set that's set up and is causing you to miss these things one after another.

DR. PARRY: Yes, right.

MEMBER BLEY: And maybe you had it earlier and it bypasses all this. If that's true, that kind of works.

DR. PARRY: I think what you're going to have to do is to -- and these questions aren't fully developed yet, but you'd have to ask the question is the additional confirmatory indication strong enough that it overcomes any previous biases. Now we can't ask it that way.

MEMBER BLEY: And that's though, when you --

(Simultaneous discussion.)

DR. PARRY: But I think we'll also find though that for the majority of standard HFEs, this is probably not going to be that big an issue.

MEMBER BLEY: Oh, I think that's right. But those aren't the ones that -- those are not the big risk factors.

DR. PARRY: They're not. Well, I'm not sure. I mean I think the -- that remains to be seen, whether they really are big risk factors.

MEMBER BLEY: You and I may disagree, but accidents tell a tale here.

DR. PARRY: Well, accidents are a little, I think are a little different though, because the -- particularly I think in the -- I think there has to be something quite significantly wrong with the procedures for them to lead into a major accident.

I mean not that the procedures are wrong, but that they're not structured for the particular scenario that you're dealing with.

DR. MOSLEH: This could be forcing him into the situation that Dennis is talking about. I think that's --

MEMBER BLEY: Right, and if all your instruments are working right and all that sort of thing. There's other stuff.

DR. PARRY: Right, okay. I mean I understand about your concerns, and it's clearly something we've got to think about. I also think that at least for an interim step in this process, it's worth looking at this type of path, because it will be a model that is more in tune with current thinking than, say, the previous ones.

MEMBER BLEY: No argument.

DR. LOIS: Dennis is not going to argue about that. I think his argument is about with respect to other questions, and the matter of the path is that got produced these questions, and that there's very little --

(Simultaneous discussion.)

DR. MOSLEH: No, I understand. We're like a style like analysis than a project, and I think your point is well taken, that it has to be a merger of the types of questions that come from this side --

MEMBER BLEY: That's right.

DR. MOSLEH: The psychology and the --

MEMBER BLEY: And experience. And looking at real experience, to make sure that from both sides, those questions are developed. I don't know if they're a different thing. Accidents are a thing we need to worry abo ut.

CHAIRMAN STETKAR: That's why we're doing this.

(Laughter.)

MEMBER BLEY: I don't disagree with that.

(Simultaneous discussion.)

CHAIRMAN STETKAR: Usually we're thinking the same thing but saying it so differently we don't recognize it. So go ahead.

DR. PARRY: So really all I wanted to do by this example, remember these are just strawmen that haven't been structured. So I want to show you the concept of that translating the PIFs that we have determined are relevant, because of understanding of a mechanism, a covenanted (ph) mechanism if you like, and using those as decision points in a tree, and then asking a set of questions which we --

And this will be a challenge, is to come up with a set of questions that are clear and can be objectively applied to help you decide which way you go on the decision tree. If we have to have multiple, you know, multiple point branches on a decision tree, that's something we might have to look at.

MEMBER RYAN: I'm reading the final conclusion, and at the end it says "yes to both questions." Let's say the answer is no to the second question, and yes to the first question. What have we got then?

DR. PARRY: I think you'd go down the branch. I'm sorry, you'd go up the branch, not down it.

MEMBER RYAN: Yes.

DR. PARRY: Yes. Remember, this is an example, okay.

MEMBER RYAN: Oh no. I'm just trying to think about the mechanics of it.

DR. PARRY: The mechanics of it would be that the set of questions would be structured such that, and we'd have to be careful about saying do I mean all of these or do I mean one of these, because that's something.

MEMBER RYAN: Okay. That's the point I'm looking for, that you have to be very careful in the language.

DR. PARRY: We do.

MEMBER RYAN: You can get mixes of yes and no's that might not give you any sort of information on where to go next.

DR. PARRY: Right, and --

MEMBER RYAN: Is that a fair point?

DR. PARRY: So that's a very fair point, I mean, because even in documents like the PRA standard that have been produced, there's some ambiguity in some of those requirements, as to whether you meet all or one of them. It's a good point, and I think that's the type of feedback we have to incorporate into this model.

MEMBER RYAN: I guess I think Dennis and John can correct me, but I would think that operator experience in those sorts of language choices would be very helpful, to really understand how operators use the language as opposed to us.

DR. PARRY: I don't think this is an operator issue as such.

MEMBER RYAN: Well they could certainly inform how that language gets used, you know. I mean do they use the same language in the same way to mean an acceptance or a denial or data, you know, or this is important versus this isn't.

DR. PARRY: Well, okay. I think that's a good point, but the questions we ask, we should also have in mind the audience to which the question is addressed.

MEMBER RYAN: Exactly.

DR. PARRY: To make sure that we're

talking about -- well, some of this is the operators. Some of this is not. Some of this is procedures, some of this is training, and some of these are plant conditions even.

So I think the question has to be phrased in such a way that the person to whom it's being asked can answer it in a positive way, or a negative way, whichever.

MEMBER RYAN: Whatever the convention is, you have to kind of line up with the conventions.

DR. PARRY: Right, right.

MEMBER BLEY: You know, I think this going to be really interesting to see where it goes. I guess one thing I'd urge, having been off to see some of the new control rooms and other things, and knowing what folks are doing in computer programs these days is if keeping it simple enough to keep it on a sheet of paper might not be as important as having a tool that lets you selectively identify things and shrink it down to get to the essence that you're after.

I'm not saying that very well, but automated tools might be necessary to make this a useful process.

DR. PARRY: I tend to be a bit of a Luddite on this side.

(Laughter.)

MEMBER BLEY: I do too. I've been doing it myself. But I think we're trying to take some fairly complicated ideas and bring them to an application point for people who probably need that kind of help, so that they don't lose track of the important things we're trying to force through this model.

DR. PARRY: Well, I think that's why embedding this into something like the HRA calculator would be a good thing to do, I think.

DR. MOSLEH: There are two aspects of what you're talking about, Dennis. One is really to have a method that captures the information in an accurate and then consistent way.

That method may be something that's too complex for an analyst actually to kind of use. You need to hide it behind a computer, an algorithm --, and therefore the idea of embedding some of these things into a computer code, to analyze, to help to aid the analyst.

A second thing is, in terms of routine applications, some of these things seem to, are going to require, you know, sequence, a series of questions that need to be asked in an organized and consistent way.

It may be difficult to fit it in the usual paper and pencil version, and the analyst might be easier guided by a computer, to go through this sequence.

(Simultaneous discussion.)

MEMBER BLEY: You're talking about prompts, different prompts depending on preceding questions.

DR. MOSLEH: So that's kind of basically, just typically use their interface, using the computer to help the analysis, rather than simplifying it to the point that you cannot sacrifice some of the key points that you want to capture.

DR. PARRY: Yes, as is somewhat accounted to that. I think we also have to recognize the fact that really we only want the model to be good enough. We don't have to have such a fine-tuning of these HEPs that we, you know, start believing them.

I mean they really are going to be pretty crude estimates. So sort of fine-tuning a crude estimate into two even more crude estimates doesn't seem to me to be all that productive.

So I think there's a balance here between doing the thing in such a way that you get as much of the relevant qualitative information out of people as possible, and force them into creating a story of what's going on at the plant and perhaps recognizing the different ways people can fail.

So that you could use it for lots of reasons, one of the reasons being is you could look for improvements for the way that -- you could look for improvements in procedures or methods of training or even plant.

CHAIRMAN STETKAR: Sorry he walked out. He'll be back.

MEMBER RYAN: It wasn't anything you said.

(Laughter.)

CHAIRMAN STETKAR: I think, Gareth, you know, that that phrase "relevant" that you tossed around a few times is really important, and I think some of the things that Dennis was talking about, and I can't speak for him, but the relevancy of some information, as far as trying to estimate human performance, may not be as clear as this simple construct, the absolute yes or absolute no.

And that's -- I don't know how to, you know. I don't have any ideas. It's just that the -- I go back to the example, and this is a simple example, but came from actual operating experience.

It's not an accident. It's just a simple example, and there are things in the literature, I'm sure, that mirror this.

A well-trained, probably the best operator I had on my shift was cooling down the plant and was turning the knob clockwise. It was cooling down too fast and he kept turning it clockwise and it kept cooling down faster and faster and faster.

He immediately knew that there was something wrong with the instrumentation controls, that something was wired backwards, and he wanted somebody to call out the instrument techs and replace this faulty thing, because we're violating technical specifications.

He didn't think that maybe he had completely the wrong mental model. He didn't even think to turn it counter-clockwise, to see if it fixed the problem.

MEMBER RYAN: Right.

CHAIRMAN STETKAR: Now he had a procedure. I'm not quite convinced that he actually ever looked at the procedure, which was a pretty simple activity. But if you asked him a lot of these questions abstractly, you would have concluded he would have never done that. He was the best guy. He would have never done that.

DR. PARRY: Yes, I know. Right. Well and I think you can't really ask operators whether they will make mistakes, because they will always tell you --

CHAIRMAN STETKAR: Well no, but even just ask them, you know, different cues that he had available and things like that.

DR. PARRY: And I think there's an element of that. Well, part of that would be that when you're looking at recovery, you're not necessarily always looking at the same crew member, right. You could be talking about somebody else.

CHAIRMAN STETKAR: Not necessarily, that's right. In fact, the crew solved the problem. The crew solved the problem by asking somebody else gee, why don't we turn this counter-clockwise and see what happens?

DR. PARRY: Yes. I think you have to build that picture of bearing in mind what's going on, and bearing in mind what failure mechanism you have in mind for that person, like the wrong mental model, okay.

Then I think you have to look for fairly convincing recovery mechanisms to make that mental model be changed. If it comes from somebody else, so be it. It's probably more difficult for somebody to alter his own mental model.

CHAIRMAN STETKAR: The only sense that I had is looking at the simple logic model we had up there which, you know, you couldn't read, but the observation was that it had all of the possible logical branches on it. Perhaps a logic model that indeed was hierarchical and had some structure to it.

DR. PARRY: Well, I think there is method in the madness of ordering the questions. They are structured to some sense logically. All the recovery stuff comes right --

CHAIRMAN STETKAR: Right.

DR. PARRY: But I think in trying to do them, I think that's going to be challenge, as like maybe putting the -- perhaps it's even senseless.

CHAIRMAN STETKAR: You're just bimodal, right. The answer is either yes or no. Is it 70 percent yes and 30 percent no.

DR. PARRY: No, no, no. There's no percentage involved at all in this, right?

DR. MOSLEH: Just looking at the combinations.

DR. PARRY: It's looking at the

combinations. So there's no probability associated with those branches. The probabilities are only associated with the HEP, given the set of conditions. But yes, it doesn't need to be bimodal.

I find it easy to write that way. But I think the, you know, my guess would be that the first questions in the tree might actually be more to do with plant conditions rather than traditional PSFs, for example.

Like you might argue one of the questions might be is this accident scenario fairly similar to another one, okay. That's probably not a very good example.

CHAIRMAN STETKAR: Oh, we know.

DR. PARRY: You know what I mean. If we're asking questions about the availability of instrumentation, for example, or the reliability of it, it would be up front, not further down, because that would certainly give you different means of failing.

DR. MOSLEH: I think, John, what we're convinced of is there needs to be a structure, some sort of a logic in our hierarchical or some other relation, maybe to many, and some way of including a non-binary kind of --

CHAIRMAN STETKAR: I don't know right at the moment.

DR. MOSLEH: But yes. We're looking at -- but I think this is the simplest of structures you can think of, as a starting point, to see okay, how do we go from the pieces to some aggregate picture.

MEMBER RYAN: Yeah, and I mean I've got two basic choices. You can either have different branching arrangements, a yes, no, a maybe, or maybe five maybes with different pathways or a whole more complicated set of yes-no questions.

DR. PARRY: Yes.

MEMBER RYAN: You want to kind of get to one or the other of those two solutions.

DR. PARRY: You know, if you're going to use expert judgment to come up with the end point probabilities, I don't think you want a million end points. You want them to be relatively small in number. So I think the argument should be that the -- have better questions as you go down the branches.

CHAIRMAN STETKAR: I think the notion of this construct is not to assign, you know, five probabilities to the various gradations, versus no or yes, because --

MEMBER RYAN: Yes. I guess to get to

that, the questions have to be very artfully created.

DR. PARRY: Right. I mean that's a big challenge. You know, I think whatever we use, ultimately, you know, Ali and Katrina are very keen on using the Bayesian belief networks, for example, and that might, you know, it certainly has the advantage of being more mathematical in some sense.

But even there, you're going to have to put in, I think, choices that are a lower level of the correlations between PIFs and things. I think in this construct, that is not -- you still talk about it, but it would be the experts that would have to come up with will be a consensus set of experts, and it would be while they're, I guess their rationale isn't so obvious, certainly the results of what they said would be.

I think if you bury it too deeply into a model, then you do get into a bit of the black box syndrome, and I mean there's -- we're always going to have a black box problem, because I think we just don't have the data to enable us to do this stuff in a more scientific way or objective way.

DR. MOSLEH: If I could add one more point to something that was brought up earlier regarding the degree of realism that we have in the branches of the decision tree, when it's set up, informed by operating experience.

In a parallel project that I'm involved, you're looking at characterizing the situational factors that were informed from operating experience, and one idea that might work is to map those to say a decision tree, a path through a decision tree, to see if you're capturing what we see in operating experience in a more explicit way.

So if you have analyzed 30 events and grouped them into like five or ten situational characterizations, do they map to the decision trees? That's sort of a sanity check regarding that.

CHAIRMAN STETKAR: Anything else from the presenters? We're a little bit over time, but this is I think a useful interchange. So I certainly don't want to cut it off because of an arbitrary clock.

(No response.)

CHAIRMAN STETKAR: No? I'd like to thank you all. I think this was really, really good. I think what I'd like to do is see --

MEMBER RYAN: Just one thing. Chris asked in the beginning, I'm sorry.

CHAIRMAN STETKAR: What I wanted to say is do you have any more focus questions on what we've seen?

MEMBER BLEY: No, I don't.

CHAIRMAN STETKAR: Okay. Thank you very much, sir. And do you sir?

MEMBER RYAN: No sir, I do not.

CHAIRMAN STETKAR: Okay, thank you very much, sir. Now what I was going to say, to answer kind of Chris' question about where do we go from here and sort of input and guidance for our next meeting, I will then ask now do you have any input on those areas, or did I cut you short too quickly, sir?

MEMBER BLEY: Well Chris actually asked us if we were, could say is this stuff right? Is it going in the right direction. Everything we've heard today I found very useful and interesting, and it sounds very good.

There are some of the stuff that holds it all together and the details are going to tell whether it's useful and appropriate to use some other times. So I don't think we're at the point yet to say yeah, this is great.

But it's encouraging and I need to see some more. I need to see some things written on the modeling aspects, the first half, and on the quantification approach. It just needs a lot of work before anybody knows how well it's going to work, I think.

MS. LUI: Yes. We're taking a phased approach. What we really want to do is to say, okay, we established the foundation, and then to get your buy-in and get your input, and if you feel that we're at the right place, then we'll continue to move forward and clearly, in the next meeting, we'll bring forward more detail.

So whatever that you feel that has been most useful for you to hear, as well as our work to support that, we will absolutely be happy to come forward with that type of information next time.

MEMBER BLEY: On the first half, on the modeling work, you know, tying all these factors together, the mid-level model as you folks call it. I think we're nearing the point that seeing something written down, to be able to study it, would be important.

On the quantification approach, I think whatever evolves between now and a few months from now would be very useful, seeing it in whatever state it's in.

CHAIRMAN STETKAR: The sense I have is I think the discipline or writing up that linkage model might also help some of the thoughts that we were discussing right here at the end, in terms of, you know, what the structure of that quantification might look like.

You know, does it need some sort of logic structure, a different type of logic structure or not. So I think I'd like -- as Dennis mentioned, from my perspective, there isn't anything that I've heard today that says, you know, change course or stop or, you know, there's something, there's a fatal flaw hidden anywhere.

I think that a key element is to -- let me just put it this way. Stop the literature search, if you will, and start documenting and provide some substance to that framework, because without that, it's really difficult for us to quite understand enough of the information to really see how it's going to work. Anything else?

MEMBER BLEY: No.

MEMBER RYAN: Yes. We talked a little bit about uncertainty and variability and those kinds of things, and I guess if I had to try and pick one word to sum all up what I'm trying to reach for, is the robustness of the model. How well does it actually describe what could be a real circumstance or set of real circumstances?

So you know, I think if you go through the exercises that Dennis and John sort of described, maybe some of that will fall out. You know, the questions are sort of -- I mean just for presentation purposes today and the detailed questions will give you some more insights about, you know, variability and uncertainty.

If, for example, you had two plants built the same way with a different operating staff, would you have exactly the same set of questions about the plants? It depends on how they were trained and what procedures they had, and how they were written versus the others.

So that's part of the fidelity, the robustness and the uncertainty aspects of any model. So believe me, it's not a calculational kind of uncertainty but more of a variability kind of concept that we're reaching for here.

But if you can think about a little bit and maybe react to that down the line, that might be interesting to hear what your insights are there.

DR. PARRY: I could give you an immediate reaction as far as the questions. I think the questions in my mind would be the same. The answers might be different.

MEMBER RYAN: Okay. What does that mean? I'm asking you to sort of --

DR. PARRY: Well, because I think the intent of the questions and the structure is to look for the variances in PIFs that would lead --

MEMBER RYAN: But the intent that you have for the question may not be the way three different respondents react to the question. You can't guarantee their response. You can only guarantee the intent of the question, not the answer.

MR. FORESTER: Is it them misinterpreting the question or having different nuances?

MEMBER RYAN: Or interpreting it in a way that you didn't think of.

MR. FORESTER: Yes, that's true. They have to be carefully written.

MEMBER RYAN: Carefully written can still be misinterpreted, and I -- well that's one person handing off to another, you know, and there's always a chance for misinterpretation. Even though it might be a small chance or a small misinterpretation, it's not impacting. Or a big interpretation change that could be impacting.

That's part of, you know, what happens if at some point in this process you get the wrong answer, wrong interpretation, miscommunication? How does that flow through the model? Does it worry, screw up your ability to interpret the model, or is it something you can live with? Does it have enough fidelity that one small misinterpretation is not going to be enough? I guess that's a measure of how good's the question. I'm trying to distinguish --

DR. PARRY: I think Ryan will like it.

MEMBER RYAN: Well, but it's the same kind of thing. I mean how -- when are you going to know the question's good enough.

CHAIRMAN STETKAR: Are you -- I want to use a simple example, that if I chose as experts five crews of operators, I mean that's sort of the thought process that I'm hearing here a bit to, you know, march through the decision tree.

You might have, depending on the interpretation of the questions and, you know, time of the day, day of the week and yadda-yadda-yadda, you might have five different sets of answers, even among the same, you know, five different crews at the same nuclear power plant.

Have you thought about that, or is this a process that's not geared toward operating crews?

DR. PARRY: I don't think that the question should be written in such a way that's an unlikely thing to happen.

CHAIRMAN STETKAR: So you wouldn't expect, you know, three crews to say yes and two crews to say no to a particular question?

DR. PARRY: It's possible. It would hope to not have that happen, but yeah, I don't know. I mean we haven't thought about the questions well enough yet.

DR. DANG: But I think the intent is not to give these questions to the crews. I mean I think we're staying with the good practices of having a team for the PRA/HRA that covers the disciplines that are needed, and that get the respective inputs to answer these questions. We're not handing over the HRA to plant experts, to --

CHAIRMAN STETKAR: Oh, is that right? I guess I'm troubled a bit by that, that you're not planning to have any input from people who understand the plant?

(Simultaneous discussion.)

DR. DANG: No. Get the input but not in the form of answering these questions directly.

CHAIRMAN STETKAR: That's a good point.

That's a good point.

DR. DANG: The questions that are placed here, of course, they make reference to the plant and the operations, but they're not meant for them to answer themselves. It's an aggregate and make a consensus.

CHAIRMAN STETKAR: That's a good point. I think the question is to help the analyst who's doing the HRA, to find out what he can from the plant people, to help him answer the question. Okay.

DR. MOSLEH: But of course if that requires that they go ask five crew and they --

CHAIRMAN STETKAR: Then you have one of these that three of them say "yes, I do" and two of them say "no, I don't." What do you do with that?

DR. PARRY: That tells you something.

DR. MOSLEH: Again, I think this is something that is difficult, in the sense of structuring a methodology and a model that is detailed enough and is prescriptive, so that we get, you know, consistent answers, as opposed to giving flexibility so that you're uncovering things that you need to know, to have a more realistic assessment of the situation.

CHAIRMAN STETKAR: And many times you

don't understand that until your replies come back.

DR. MOSLEH: So that balance, that optimization is something that we are not -- although we have explored many different styles of analysis, and we have a team here that has spent a lot of time to see how we can structure the analysis, such that we get consistent results with a certain level of detail that would be at that level of consistency. But these things we need to kind of try ourselves, before we can share with you.

MEMBER RYAN: Thank you. That's critical.

CHAIRMAN STETKAR: One last thing. I guess we ought to -- it's probably premature to think about when we'll meet again, but given the fact that you've sort of targeted getting things out on the street by about a year from now, let's say, we should certainly, in the introduction, I think Erasmia said quarterly meetings, our particular constraints on subcommittee meetings and things like that.

I think we should probably target for, you know, March, no later than April time frame. That's not quite a quarter, but since we're -- we generally don't meet in January, and even if we do, our January schedule is already full, and we're out into February already. We're looking at probably early March is the earliest. But I think we ought to reconvene some time in that time frame.

DR. LOIS: And by that time, you would like to see some of the --

CHAIRMAN STETKAR: We would certainly, I think I would say that we should not reconvene and perhaps my colleagues here could correct me, until we have a reasonably coherent documentation of that mid-level model, if that's what we're calling it. I think that more sort of philosophical discussions, if I can call it that, about that, would not necessarily be all that productive.

So I think you need to look at when that might be available. It doesn't have to be polished. I mean these are not final committee meetings. But it ought to have some prose attached to it and rationale. I think whatever else on the quantification part of it might fall out.

But I think without that basic documentation that we can actually have, and remember, what I'd like to do is have that available traditionally a month, you know, 30 days before our subcommittee meeting, so that we actually have a chance to digest it. That may dictate, you know, when you need to come back to us. You can just coordinate with John as far as setting the schedule.

The message is think about it sooner than later, because our schedules these days are miserable. I'll just throw that out.

MEMBER RYAN: No they're not, John. They're full of fun and wonderful things to do, all day, every day.

CHAIRMAN STETKAR: They are. That too, and with that, unless Christina is -- before we miss, is there anybody, members of the public? I don't see anyone here, but I always have to ask for that, that wants to say anything?

(No response.)

CHAIRMAN STETKAR: With that, Christina?

MS. LUI: Well yes. We certainly appreciate you spending time with us this afternoon, and we would aim to get back to you by March time frame. I know my project team may not agree with me, but that's my issue.

CHAIRMAN STETKAR: One message is we probably can't -- we can't get a meeting in before March. I'd like to get one in kind of in the March to April time frame, if you can support it, so that we don't drag it too far.

MS. LUI: Yes absolutely, because I would like to probably schedule another meeting prior to any documentation goes out on the street. So March time frame would be the right time frame before we actually would be able to squeeze in another interaction with you before the final set of documentations.

I do have one question for you guys. I understand that you'll be meeting with the Commission next month.

CHAIRMAN STETKAR: Yes.

MS. LUI: Do you need anything from us to support any type of discussion in this area?

CHAIRMAN STETKAR: No.

MEMBER RYAN: I don't think that's on our agenda.

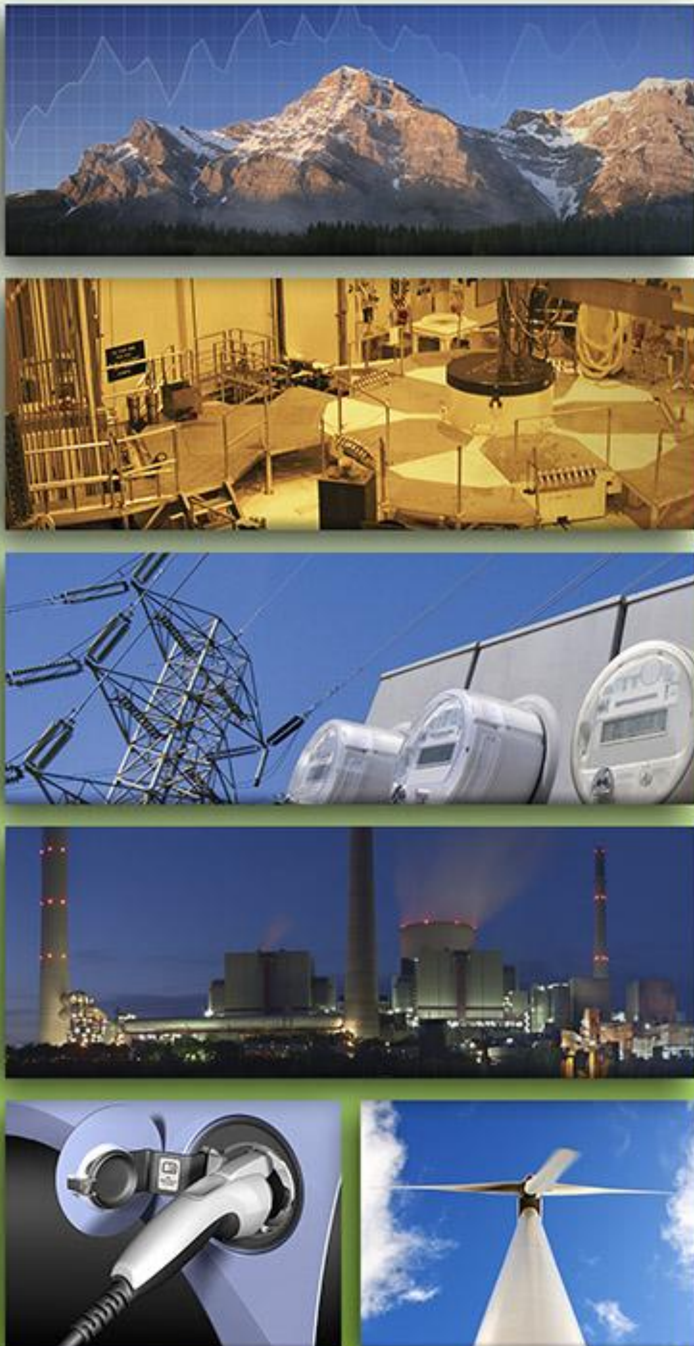
CHAIRMAN STETKAR: No. Nothing on our agenda in this area at all.

MS. LUI: Okay, all right.

MEMBER RYAN: And the questions and answers, we may get a question from at least one of the Commissioners.

CHAIRMAN STETKAR: We might, but probably not. I'm betting no. Given the other topics on the agenda, I'm betting no. Anyway, with that, we're adjourned.

(Whereupon, at 5:29 p.m., the meeting was adjourned.)



ELECTRIC POWER RESEARCH INSTITUTE

Development of a Hybrid Model for Human Reliability Analysis

EPRI Perspective

Stuart Lewis
Program Manager
Risk and Safety Management

ACRS PRA Subcommittee Meeting
October 18, 2010

Status of HRA Research at EPRI

- Basic HRA methods developed by EPRI in late 1980s/early 1990s still in wide use within nuclear industry
  - SHARP1 framework for HRA (including guidance for qualitative analysis)
  - Complementary methods for representation and quantification of human failure events
    - HCR/ORE correlation for time-constrained human actions
    - Cause-based decision tree to capture influences not directly accounted for in a time-reliability correlation

Role of the HRA Calculator®

- Focus for past ~10 years has been on development of the HRA Calculator®, a software tool
- Objectives of the tool
  - Facilitate performance of human reliability analysis
  - Promote consistency among different analysts addressing the same problem
- Tool implements methods for
  - Assessment of pre-initiator human failure events (HFEs): THERP, ASEP
  - Assessment of post-initiator HFEs: HCR/ORE and CBDT for cognitive failure, THERP for execution
  - Evaluation of dependencies among post-initiator HFEs

Status of the EPRI Methods/Tools

- HRA Calculator® is now used by every nuclear utility in the US
- Methods are generally well-understood by users, but
- There is recognition that the methods
  - Are aging, without significant review or update for 18+ years
  - Included elements that were meant to be examined further over time

Motivation for EPRI involvement in hybrid model development

- Take advantage of work done by NRC researchers to provide
  - Better psychological underpinning to HRA
  - More comprehensive understanding of potential human failure events
  - Updated approach to quantification

Expectations for the hybrid model

- Practical to implement
- Not requiring major restructuring of current PRA models in the near- to mid-term
- Achieving consistency and repeatability in results

Together…Shaping the Future of Electricity

Addressing SRM-M061020 on Human Reliability Analysis Model Differences

Erasmia Lois, PhD
Senior Risk and Reliability Analyst
Division of Risk Analysis
Office of Nuclear Regulatory Research

ACRS PRA Subcommittee Meeting

October 18, 2010

Outline

- Background
- Interpretation of the SRM
- Approach and Aim
- Focus
- Interactions with the ACRS
- Objective of Today's Meeting
- Who is Involved in the Project
- Challenges
- Schedule
- Anticipated Uses

Background

- SRM-M061020 directed the ACRS to "work with the staff and external stakeholders to evaluate the different human reliability models in an effort to propose a single model for the agency to use or guidance on which model(s) should be used in specific circumstances"
- Through interactions with the ACRS, RES initiated work to support addressing the SRM
- ACRS input into the work
- Collaborative with EPRI

Interpretation of SRM-061020

- Interpret "model" in the more general sense of "method"
- A single method is the most desirable
- Use of more than one method should be justified
  - Why more than one
  - Which methods should be used for which applications
  - Need for implementation guidance for each particular method and application
- Desirability for convergence of NRC and industry on HRA methods

Approach and Aim

- Establish a consensus approach by developing a “single” high-level method/structure that ensures
  - consistency through out the analysis process
  - be of sufficient generality to support application for different domains
    - Applications such as low power and shutdown, ex-control room actions, external hazards, and Level 2 analysis may require adaptation of the overall structure for addressing domain-specific needs
- Gain acceptance from PRA/HRA and human factors experts and practitioners

Focus

- Current focus
  - Start with addressing the issue for a detailed internal event/at power PRA/HRA
  - Converge with EPRI on HRA methods(s) and practices
- Future
  - Expand to other scenarios
  - Modify for screening/scoping analyses

Interactions with the ACRS

- Subcommittee briefing, April 7, 2010 during which we presented a "hybrid" approach and ACRS members posed many questions summarized as follows
  1. Why not choose an existing method
  2. Is it a "hybrid"
  3. How the needs of different domains (e.g. LPSD) will be handled
  4. Would the approach be suitable for regulatory applications using existing PRAs
     - What could be the impact on existing PRAs
  5. Would it be suitable for new PRAs
  6. Need to understand the various facets of the approach
  7. How quantification will be handled
  8. How will obtain user buy-in
- Current plan for quarterly meetings with the Subcommittee
  - Need for ACRS members develop an in-depth understanding of the technical work
  - Different aspects of the work to be discussed in different workshops
  - 1st workshop today, Oct.18, 2010

Objective of Today's Meeting

- Discuss items 1-5, showing the rationale lead to the proposed approach
- Discuss part of question 6 related to building a technical basis for HRA using results and inputs from cognitive research and expertise—what we called before mid-layer model
- Present an overview of the current thinking to address quantification—bullet 7
  - Recognizing that quantification is the area that has many facets needed to be addressed we are going forward with short term and longer term activities.
    - Only the short term perspectives will be discussed today
    - The discussion on quantification shows early stages of the work
- Obtain feedback and input on the work being performed
- Plan for the next meetings
  - Topics to be discussed
  - Schedule

Who is Involved

- Collaborative with EPRI
- RES staff
- NRC-sponsored organizations
  - Sandia National Laboratories
  - Idaho National laboratory
  - University of Maryland
  - Paul Scherrer Institute, Switzerland
- Interdisciplinary Expertise
  - PRA/HRA
  - NPP Operations
  - Human factors/cognitive psychology experts

Challenges

- High-level concepts appealing, resolving issues showing up in the details
- Effectively communicating view points and achieving consensus among the project team, comprised recognized experts in different disciplines
- ACRS buy-in
- Facilitating understanding and acceptance by the larger community, both in the NRC and within the industry

Schedule

- Quarterly working meetings with the ACRS subcommittee
  - First scheduled for October 18
- Technical Basis for public review—Sept 2011
- User's Guide—Sept 2011
- Application on event evaluation for public review—Sept 2011
- Final report—2012
- Computerized capability—TBD

Anticipated Uses

- The tools developed should support NRC staff risk-informed activities
  - Event evaluations (significant risk determination)
  - Guidance for staff review of risk-informed licensee requests
    - Current plants—e.g., risk-informed licensing changes
    - Future plants—reviews of PRAs/HRAs for new plant operation licensing and practices
  - New PRAs and especially the Level 3 PRA planned by the staff for the next few years
- Industry applications in the same areas

Basis for Use of a "Hybrid" HRA Approach to Meet SRM-M061020

ACRS PRA Sub-Committee Meeting

October 18, 2010

Presented by

John Forester (SNL) and Vinh Dang (PSI)

Objective

- Inform on the approach taken to address SRM-M061020

Presentation Outline

- Interpretation of SRM and initial process
- Results of evaluation of HRA Methods
- Implications for SRM
- Current thinking on a "hybrid" approach

Interpretation of SRM-061020

- A single method is the most desirable
- Desirability for convergence of NRC and industry on HRA methods
- Use of more than one method should be justified
  - Why more than one
  - Which methods should be used for which applications
  - Need for implementation guidance for each particular method and application

Initial Process for Addressing SRM-M0601020

- Survey of NRC staff to identify regulatory applications in which HRA plays an important role and what issues the staff is dealing with when applying HRA
- Workshops with NRC, national laboratory and EPRI staff to obtain their views on the path forward (April and June 2009)
- The need to build on existing methods and experience was emphasized

From the Workshop - Review Methods Against Evaluation Criteria

- **Validity – content, construct**
  - *models*
- **Empirical validity and quantitative performance**
  - *quality of numbers*
- **Reliability, consistency, traceability**
  - *variability*
- **Usability and resources—**
  - *Having a **practical** tool is key*
- **Scope of applicability**
  - *From scoping to detailed HRA*
  - *For different PSA scopes*
- **Experiences with method; installed base (users and applications)**

Inputs to HRA Method Evaluation

- NUREG-1792, Good Practices
- NUREG-1842, Evaluation of HRA Methods vs. Good Practices
  - Theoretical/scientific considerations - examined their strengths, weaknesses, and underlying basis
- Experiences from NRC and other reviews of HRAs
  - Applications experience/practicality
- International HRA Empirical Study (additional methods addressed)
  - Performance-based HRA method evaluation

The International HRA Empirical Study (2006-2010)

- **Simulator study at the Halden Reactor Project**
  - 12-organizations; 13 methods; 14 crews
  - Assess strengths and weaknesses of HRA methods based on predictive analyses vs. crew data
  - Evaluate and identify sources of variability in results
- **Major insights from the Empirical study**
  - Significant variability in results across HRA methods and within teams for one method
  - Need for more comprehensive/consistent qualitative analysis – to allow consistent identification of correct issues
  - Ability to incorporate the issues into quantification

Evaluation of HRA Methods

Findings across methods

- **On evaluation, all methods have some significant shortcomings and different strengths**
  - Relates in part to originally intended scope of methods
  - Relates to trade-offs
    - e.g. simplicity vs. ability to represent broad range of HFEs and performance conditions,
    - e.g. repeatibility vs. ability of analysts to consider specific influences
- **Qualitative analysis is a shared weakness -** some methods much better than others

  (scope - issues and influences considered, guidance for analysts, translation into quantification inputs)
  - **Impacts the consistency** of estimated HEPs
  - **Impacts validity of HEPs**, to the limited extent that this validity can be empirically verified

Findings across methods (cont.)

- **Two common practices in HRA are**
  - **Mixing and matching of methods**
    - e.g. decision/cognition vs. implementation/execution
    - e.g. different methods for full-power vs. shutdown PSA
    - Motivation is to allow consideration of different performance influences
  - **“Adjustments” of method relative to reference method guidance**
    - Variation in how formalized the adjustments are
      - From a documented set of rules to apply to analyst discretion

These practices also affect consistency of estimated HEPs and their validity.

Evaluation of selected method features

- **Time Reliability Curves for decision/diagnosis**
  - easy to use
  - difficult to adjust for other influence factors
- **Prescriptive rules for analysis, including worksheets**
  - supports repeatability, traceability, and can support validity
  - need to develop different sets for each context, e.g. full-power vs. shutdown PSA
- **Binary quantification inputs**, e.g. yes/no
  - may support consistency
  - "cliff-edge" effect of assigned inputs
- **Multiple-level quantification inputs**, e.g. L-M-H, 0-10
  - allows more nuanced factors
  - scaling guidance (anchors) needed to support repeatability, traceability, inter-analyst reliability

Evaluation of selected method features

- **Narrative-based failure mechanisms**
  - supports broad set of influences, plant- and scenario-specific error mechanisms
  - requires expertise in plant-specific operations
  - difficult to structure analysis, to review
  - effort and expertise in elicitation

HRA Method Evaluations - Implications for SRM to Address Criteria
(e.g., validity, consistency, traceability)

- **Extend guidance for qualitative analysis**
  - Important to address a comprehensive set of performance issues and to obtain consistent quantification inputs
- **Support for analysis and representation of failure mechanisms in some detail**
  - As in ATHEANA, MERMOS, and other methods that explicitly link performance conditions to potential behaviors and failures
- **PSF rating guidance to support comprehensive analyses, repeatable and traceable results**
  - Consistent inputs to quantification, i.e. *qualitative-quantitative interface,* once key performance issues are identified
- **Improve technical basis by linking the HRA to accepted models of human performance**
  - *In addition to consideration of experience and engineering judgment*
  - Importance for acceptance of HRA methods and their results

"Hybrid" Approach

Capitalize on Strengths of Existing Methods and Psychological Models/Data

- The hybrid is (<u>currently</u>) building on
  - Information from cognitive psychology to build an explicit connection (human response model) between:
    - the human failures leading to a human failure event (HFE) (proximate cause)
    - the underlying psychological failure mechanisms
    - the factors driving the failure mechanisms (PSFs and plant/scenario conditions)
      - **Improves validity** – helps identify what we need to measure and what specifically we need to predict
  - ATHEANA concepts for identifying and incorporating contextual aspects associated with human performance
    - Needs more structure for capturing relevant context and may need to be extended to be more comprehensive
    - **Improve validity and consistency**

"Hybrid" Approach
(Continued)

- The hybrid is (<u>currently</u>) building on:
  - The causal structure of the Cause Based Decision Tree (CBDT) method for quantification or some other structured approach (e.g., Bayesian Belief Networks)
    - Key is to have a concrete, **structured** approach for quantification that capitalizes on psychological information and incorporates critical drivers of performance
      - Provides clear guidance for what information needs to be collected, what decisions need to be made, and how to make those decisions
    - **Improves validity, consistency, transparency, traceability**
    - **Requires less analyst expertise and fewer resources**
  - Use of a PRA-type explicit structure (CRTs) to support identification of potential failure paths and associated context for HFEs
    - **Improves validity and consistency**

“Hybrid” Approach

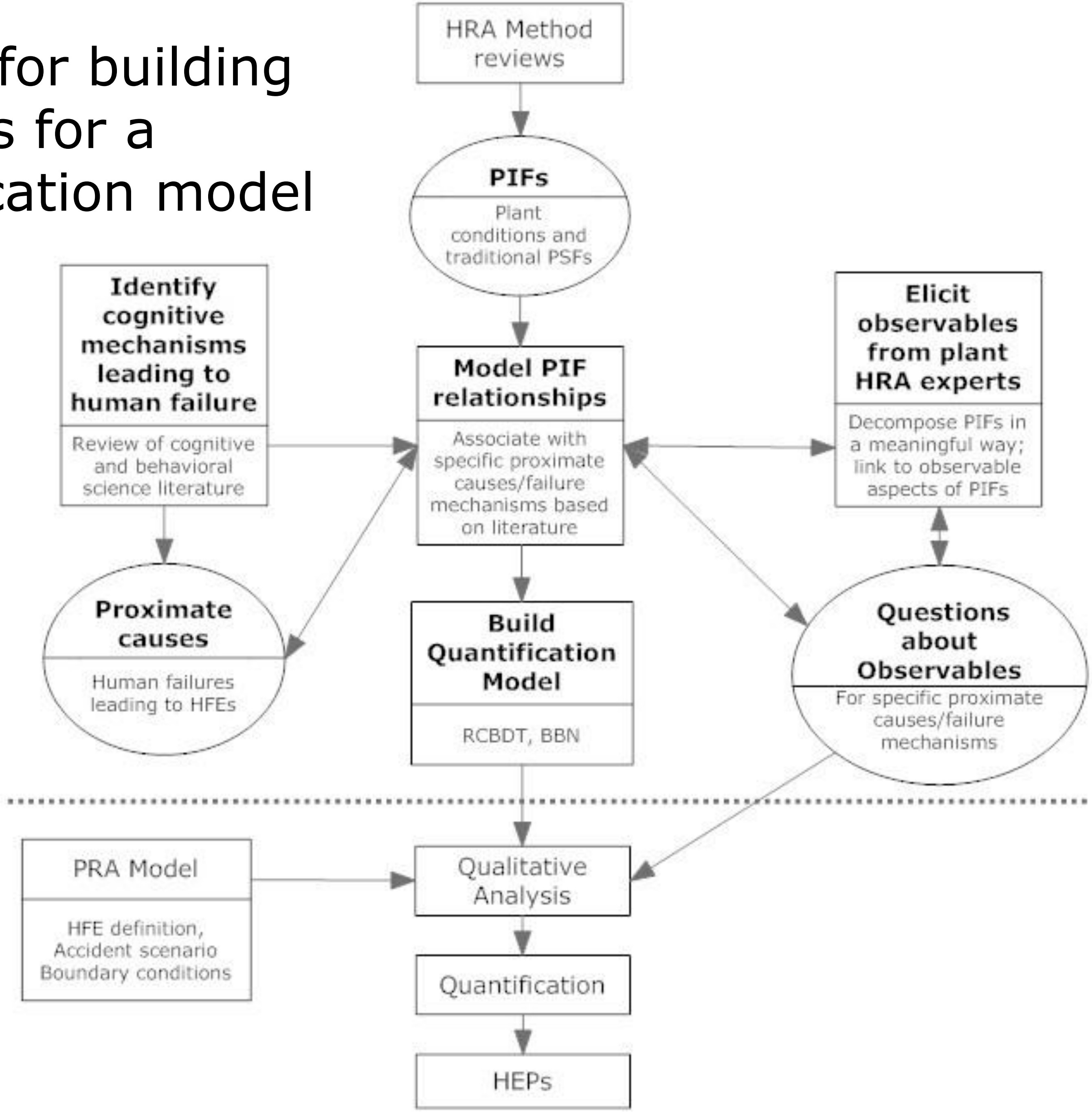
- Incorporating in the method an explicit process for capturing current understanding of human behavior from both the literature, NPP operations, and PRA/HRA experience
- The method should provide a structure to guide the analysts to ask/identify the right issues during the information collection and qualitative analysis task and at the same time is:
  - Improved scientific basis
  - Structure applicable cross domains
  - Consistent with industry (another driver supporting use of CBDT type structure)
  - **Practical**

Next Presentations

- Human performance model from psychological models/data (formerly mid-layer model) to support HRA
- Experimentation with use of decision tree approach for quantification

Back-Up

Process for building the basis for a quantification model



Key Strengths & Weaknesses (1)

	Strengths	Weaknesses
THERP	designed to **address large set of NPP tasks** included in PSAs several **validations of execution models** **large installed base**	limited **treatment of cognitive demands after event identification and in execution** **basis for diagnosis model** (for TRCs) **very few PSFs** and **limited guidance** for their assessment **no review of execution model** with respect to **modern HMIs**
ASEP	**more specific rules for analysts** (relative to THERP) **ease of use**	**same weaknesses as THERP** due to common basis **more limited than THERP** in modeling failure mechanisms (proximate causes), but can in principle be combined with THERP as needed
CBDT	specifies **failure mechanisms and their relation to plant and PSF context** **explicit set of causal factors and failure mechanisms** binary assessment of factors **avoids need for fine-grained judgments**, advantage for consistency **ease of use**	basis of the **decision tree branch HEPs** **limited set of causal factors** and failure mechanisms **binary assessments of causal factors** (can be modified but introduces more branches)

Key Strengths & Weaknesses (2)

	Strengths	Weaknesses
SPAR-H	highly **traceable worksheets** **scaling guidance** for rating of PSFs **ease of use**	**ambiguous and inadequate** guidance for some PSFs **limited set of factors** and influences considered limited ability to **model specific (narrative-based) failure mechanisms** limited ability to consider **cognitive demands of execution**
HCR/ORE TRCs	**curves derived from simulator data** (but not validated)	**no validation of TRCs** **no limit on relation of response time to available time** (accounting for magnitude of time window) lack of support for **assumption that slow performance is poor performance** (may not credit crews' deliberate use of available time) in practice, **median response times are frequently estimated without simulator data** (not based on plant-specific and scenario-specific simulator data)
FLIM	**PSF scaling guidance for 7 key PSFs,** relating PSFs to failure mechanisms **comprehensive and broad definitions of PSFs** **structured elicitation of NPP operator inputs**	need for **external calibration** values **PSFs may compensate each other unrealistically** (additive model of PSF effects) / underrepresent of impact of very poor individual PSFs **other quantification issues** (grouping for calibration, updating of HRA) **resources** for elicitation, impact **on PSA model development and updates**

Key Strengths & Weaknesses (3)

	Strengths	Weaknesses
ATHEANA	encourages consideration of **broad set of influences and failure mechanisms** qualitative analysis **guidance more extensive than many methods**, based on psychological models of cognition and human performance in complex domains encourages **holistic view of influences and their interrelationships**, including dependences **evaluates a range of conditions**, more explicit consideration of **aleatory influences** uses **plant-specific operations information** elicited from **plant staff**	**"open" set of factors** and failure mechanisms requires **dedicated expertise for plant staff elicitation** reliance on **expertise on integration of influences** to consider in qualitative analysis (and on interview approaches) **external review may be hindered** by reliance on plant-specific expertise **repeatibility** of qualitative and quantitative analysis (may call for more structure) **expert judgment in quantification** level of **resources and required expertise**
MERMOS	identification and evaluation of **multiple, detailed failure mechanisms (narratives)** evaluates a **range of conditions**, more explicit consideration of **aleatory influences** **flexible but consistent structure** for failure narratives uses **plant-specific operations expertise** accounts for **observed performance in NPPs** and **structure of NPP operators' task** quantification uses **context features that are**	**"open" set of factors** and failure mechanisms **reliance on operations expertise** **difficult to review plausibility / completeness of failure narratives** without specific expertise concerning on plant operators and how they work **expert judgment in quantification** level of **resources** and required expertise

Method Evaluation Criteria – details (1)

- **Validity – content, construct**
  - Addresses performance issues and factors/influences accepted to be important
  - Based on an accepted model of human performance, relevant to NPP emergency operations
- **Empirical validity and quantitative performance**
  - Bases for the method's values (tabled values, nominal HEPs)
  - HEPs estimated with the method are consistent with data and operating experience
- **Reliability**
  - Consistency among analysts and across applications by an analyst
  - Reproducibility
  - Traceability (qualitative and quantitative)

Method Evaluation Criteria – details (2)

- **Usability and resources**
- **Scope of applicability**
  - Support for scoping, coarse/fine screening, and detailed analyses
  - Extensibility beyond full-power internal events PSA (ability to incorporate elements to address performance issues and factors arising in other PSA scopes)
- **Insights from practical experience with method (applications, HRA reviews), installed base**

ASEP Comparison Against Criteria (EXAMPLE)

Criterion	Positive	Negative
Content Validity	• Cognition and Action covered • Dependency between HFEs covered through THERP • EOO included • Limited coverage of recovery • Accounts for uncertainty	• Minimal coverage of context • Cognition can be ignored • No PSF interaction • Minimal inclusion of EOC • No coverage of failure mechanisms • Minimal task decomposition
Empirical Validity	Origin is THERP	
Reliability		Inter-analyst variability in level of modeling detail in specifying HFE and in how each PSF is assessed
Traceability	Mathematical tracing possible	Qualitative tracing possible, but dependent on level of analysis and documentation
Level of Analysis	Screening and scoping level	No detailed level
Usability	Fairly straightforward application	

HRA Methods Evaluated Against Criteria

ATHEANA	A Technique for Human Event Analysis
CBDT*	Cause-Based Decision Tree
SPAR-H	Standardized Plant Analysis Risk - HRA
THERP	Technique for Human Error Rate Prediction
ASEP	Accident Sequence Evaluation Program

*CBDT was evaluated with the understanding that it addresses only the cognitive failure probability

ATHEANA Comparison Against Criteria

Criterion	Positive	Negative
Content Validity	Fairly complete qualitative analysis – covers all elements	
Empirical Validity	HEPs based on expert judgment – could be seen as positive or negative	
Reliability		Mathematical reproducibility difficult
Traceability	Qualitative traceability possible (dependent on documentation by analyst)	Mathematical traceability difficult
Level of Analysis		Does not offer a screening or scoping option
Usability		Time and resource intensive

CBDT Comparison Against Criteria

Criterion	Positive	Negative
Content Validity	• Plant and PSF context • Cognition and Action covered • EOO included • Failure mechanisms explained	• No crew context • No PSF interaction • Partial dependency between HFEs • No EOCs • Partial coverage of recovery • No quantification of uncertainty • No guidance on task decomposition
Empirical Validity	Based on THERP and adapted by expert judgment	
Reliability	Quantitative analysis reproducible as long as qualitative analysis is same	May be significant differences in the qualitative analyses between analysts
Traceability	Mathematical traceability possible	Qualitative tracing possible, but dependent on level of analysis and documentation
Level of Analysis		Does not offer a screening or scoping option
Usability	Fairly straightforward application and fairly easy to use within HRA Calculator	

SPAR-H Comparison Against Criteria

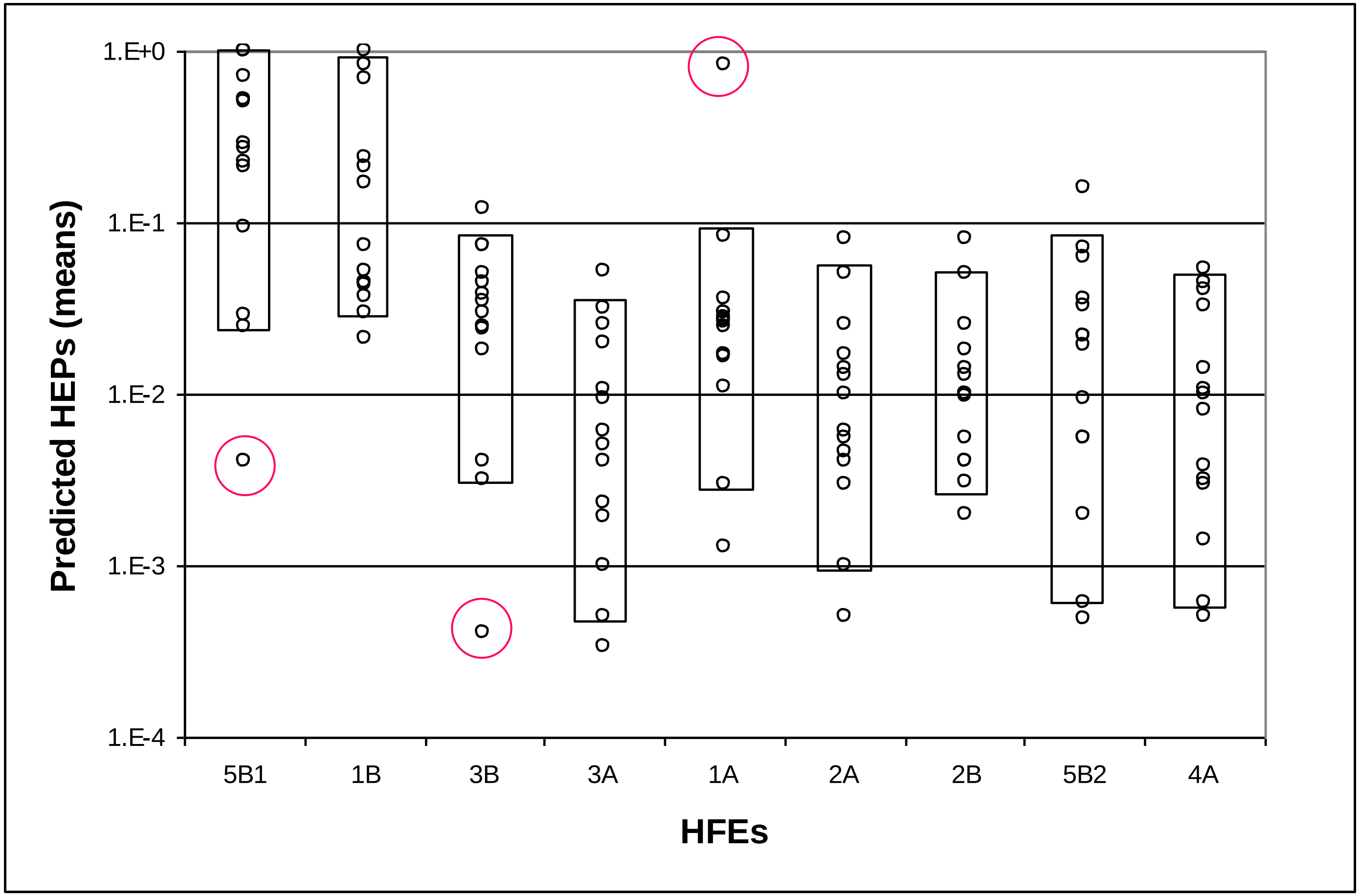
Criterion	Positive	Negative
Content Validity	• Cognition and Action covered • Dependency between HFEs covered • EOO included • Failure mechanisms explained • Accounts for uncertainty	• Minimal coverage of context • No PSF interaction • Minimal inclusion of EOC • Partial coverage of recovery • No coverage of failure mechanisms • No task decomposition
Empirical Validity	Compared multipliers uses to assess effect of PSFs to other PSF-intensive methods	
Reliability		• No guidance offered on checking reasonableness of final HEP • Possible differences in selection of important PSFs and impact of PSFs
Traceability	Mathematical tracing possible	Qualitative tracing possible, but dependent on level of analysis and documentation
Level of Analysis	Somewhat scoping and detailed	No screening option
Usability	Easy to use	

THERP Comparison Against Criteria

Criterion	Positive	Negative
Content Validity	• Cognition and Action covered • Dependency between HFEs covered • EOO included • Recovery included • Accounts for uncertainty	• Minimal coverage of context • No PSF interaction • Partial coverage of HFE dependency • Minimal inclusion of EOC • No coverage of failure mechanisms • Partial task decomposition
Empirical Validity	Mix of empirical data and expert judgment	
Reliability		Reproducibility reliant on task analysis done – variability possible in level of detail of task analysis
Traceability	Mathematical tracing possible	Qualitative tracing possible, but dependent on level of analysis and documentation
Level of Analysis	Screening and detailed level	
Usability	Screening level fairly easy to apply	Detailed level more resource intensive

Range of predicted mean HEPs - SGTR

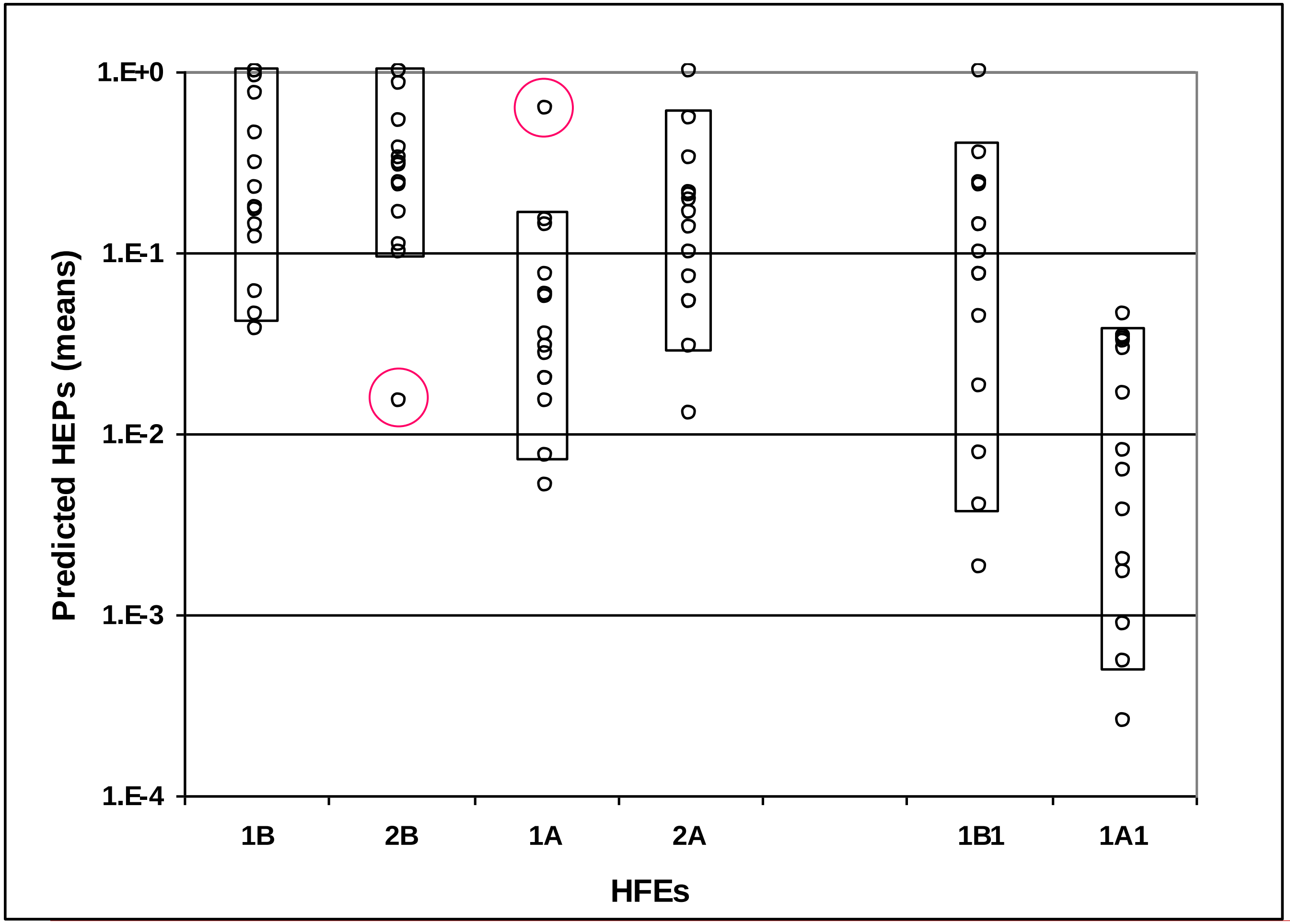
Boxes drawn around range, 1 maximum value and 1 minimum value excluded from each range.



- After exclusion, most ranges span < 2 orders of magnitude
- Many outliers relatively close to the range.
- Exceptions are highlighted.
- Exceptions explained in terms of inconsistent assumption, and method application

Range of predicted mean HEPs - LOFW

Boxes drawn around range, 1 maximum value
and 1 minimum value excluded from each range.



- After exclusion, ranges span < 2 orders of magnitude
- Many outliers relatively close to the range.
- Two exceptions are highlighted.
- These are being examined to determine causes
  - method or assumption or combination

Establishing a Technical Basis for Crew Performance Model in HRA

Ali Mosleh

ACRS PRA Sub-Committee Meeting

Oct 18, 2010

Team:

April Whaley, Ron Boring, Johanna Oxstrand, Dana Kelly (INL)

Stacey Hendrickson, John Forester, Katrina Groth (SNL)

Song-Hua Shen, Jing Xing, James Chang (NRC)

Ali Mosleh (UMD)

Outline

- **Background and Goals**
  - Establishing a Technical Basis for HRA Crew Performance Model
- **Simple Causal Perspective**
  - Guiding the psychological literature search
- **Literature Search and Scope**
- **Findings and Examples**
- **Current Status**

Background and Goals

- Addressing issues related to “Content Validity” is needed to address variability
- Content validity depends on the theoretical frameworks and models utilized by methods and the extent to which they reflect the current understanding of human performance
  - Need to incorporate findings from current psychology and operations experience
- Content validity also depends on an accurate and consistent use of the frameworks and models employed
  - Current practices depend on analyst expertise for content validity
  - Need to formalize the process to reduce dependency on analyst expertise and expert judgment

Background and Goals

- This effort aims to provide a cognitive framework that establishes links between PRA scenarios, context, operator psychological processes, and resulting performance
  - Supported by psychological literature
  - Informed by operating experience
- Further improve content validity through formalization
- Address practicality and consistency through tools in which a sound technical basis underlies the tool

Background and Goals

- Relation to Evaluation Criteria:

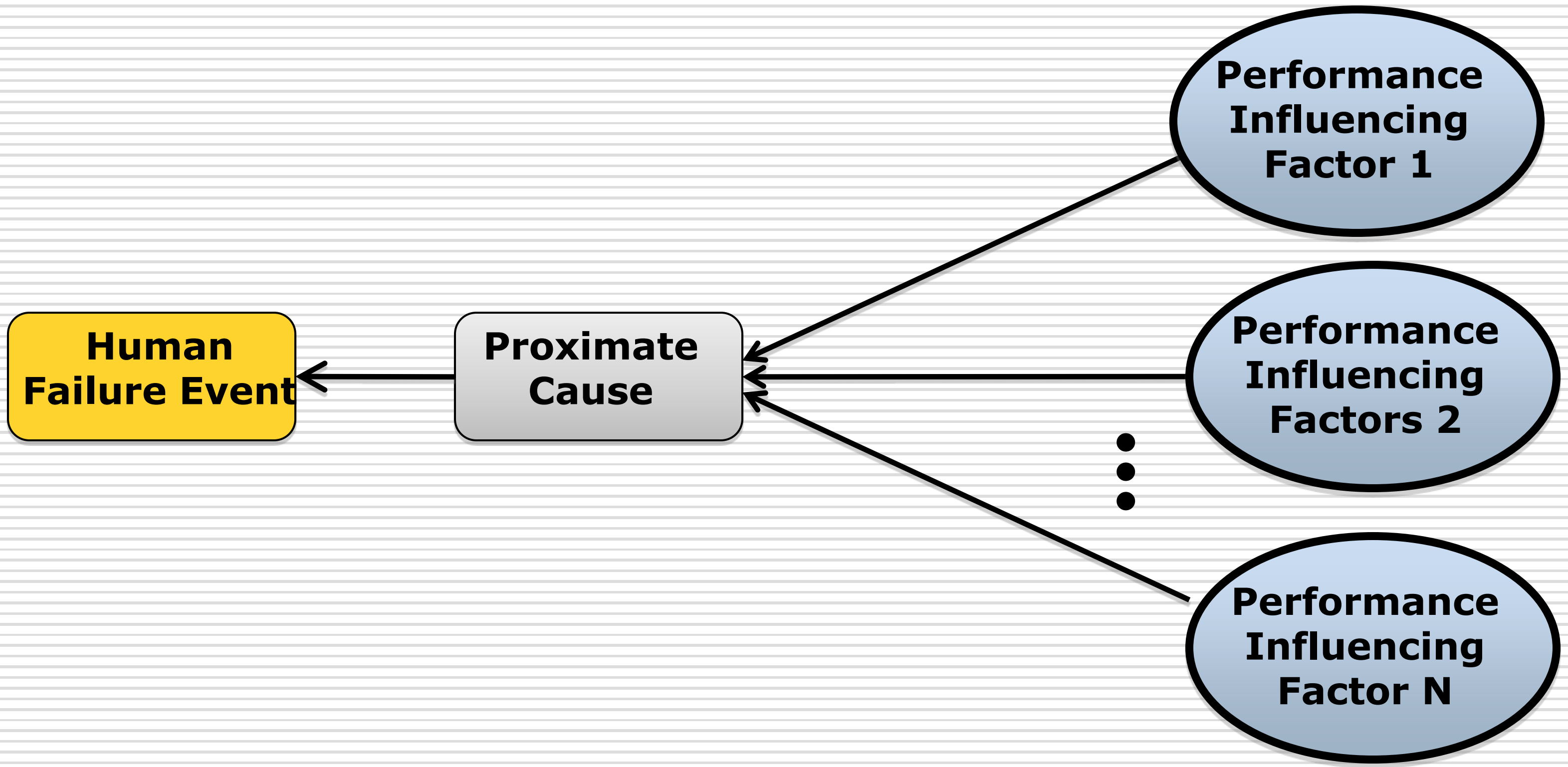
  Validity – content, construct

  Models

  A causal perspective with empirical or theoretical basis for its constructs (elements and relations) is a key scientific approach to creating valid predictive or descriptive models

- Products of the effort
  - Practical Model for Use in HRA
  - Technical Basis (theory, literature support, …)

A Simple Causal Perspective for Guiding the Psychological Literature Search



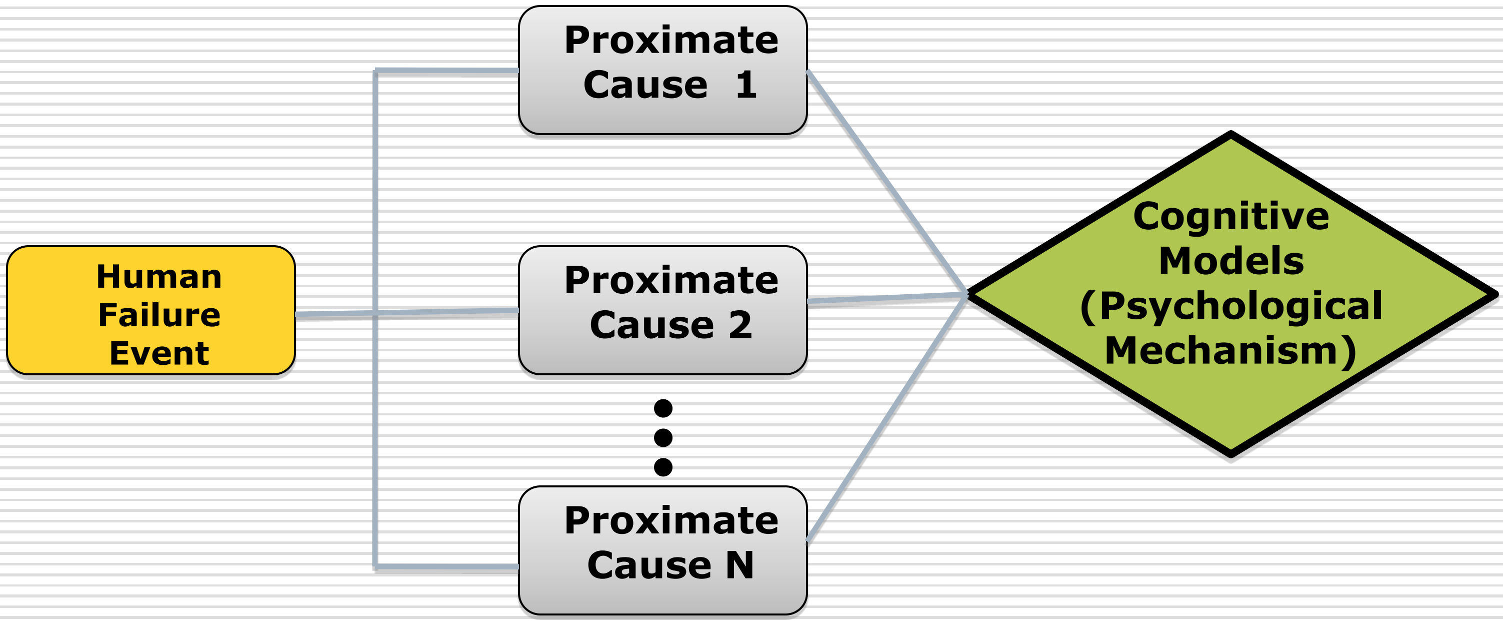
Context Factors (PIFs)

- Plant
  - Status (past and current) of
    - Critical safety functions
    - Relevant systems/components/equipment
  - Dynamics
    - Rate of change/speed of event evolution
    - Time
- Crew Mental model
    - Where the crew thinks the plant is headed
    - Crew understanding of the event
  - Goals
    - Related to the plant and event, both general and event-specific
    - Organizational goals
    - Action history
- Human-Machine interface/ergonomics, training, procedures, and other typical PSFs

Psychological Literature Review

- Reviewed relevant psychological, human factors, and human performance literature
  - Past 10 years of *Annual Review of Psychology*
  - The 50th Anniversary Special Edition of *Human Factors*
  - Cognitive Psychology and Human Factors/Human Performance textbooks
  - Related peer-reviewed articles and sources
- Methodology
  - Identify appropriate models based on
    - Degree of community support/consensus
    - Relevance to NPP operating environment
  - Fill Gaps, interpret, extend, generalize
  - Synthesize and Simplify

Search for Proximate Causes in Psychological Literature



Psychological Literature Review

- Sample Models, Frameworks, and Theories Reviewed
  - Sensation and Perception, Information Foraging Theory, Situation Awareness and Sensemaking
  - Recognition-Primed Decision Making, Situation Awareness and Sensemaking, Aided Decision Making, Cognitive Biases, IDAC, Cognitive Engineering
  - Human Performance Models, Contextual Errors, Omission-Provoking Factors (e.g., James Reason's work on slips, lapses, and maintenance errors)

Use of Psychological Literature for Proximate Causes, Mechanisms, and PIFs

- Criteria used was that to the extent possible Proximate Causes should
  - have distinct non-overlapping definitions
  - be observable or inferable in a practical manner

- Search approach
  - PCs, Psych Mechanisms and PIFs identified by various theories and models, absent any top-down structure

Use of Psychological Literature for Proximate Causes, Mechanisms, and PIFs

- Complementary approach applied
  - Using "synthesis models" or extrapolation from main psychological model adopted
  - Applying a top-down structure of relevant dimensions in NPP context
    - Example: For information gathering dimensions included were
      - Information Source (plant, procedure …)
      - Information Modality (Auditory, Visual, …)
      - Response Mode (Active vs Passive information gathering)
- Proximate cause and psychological mechanisms were defined in more general terms to allow application over wider range of situations

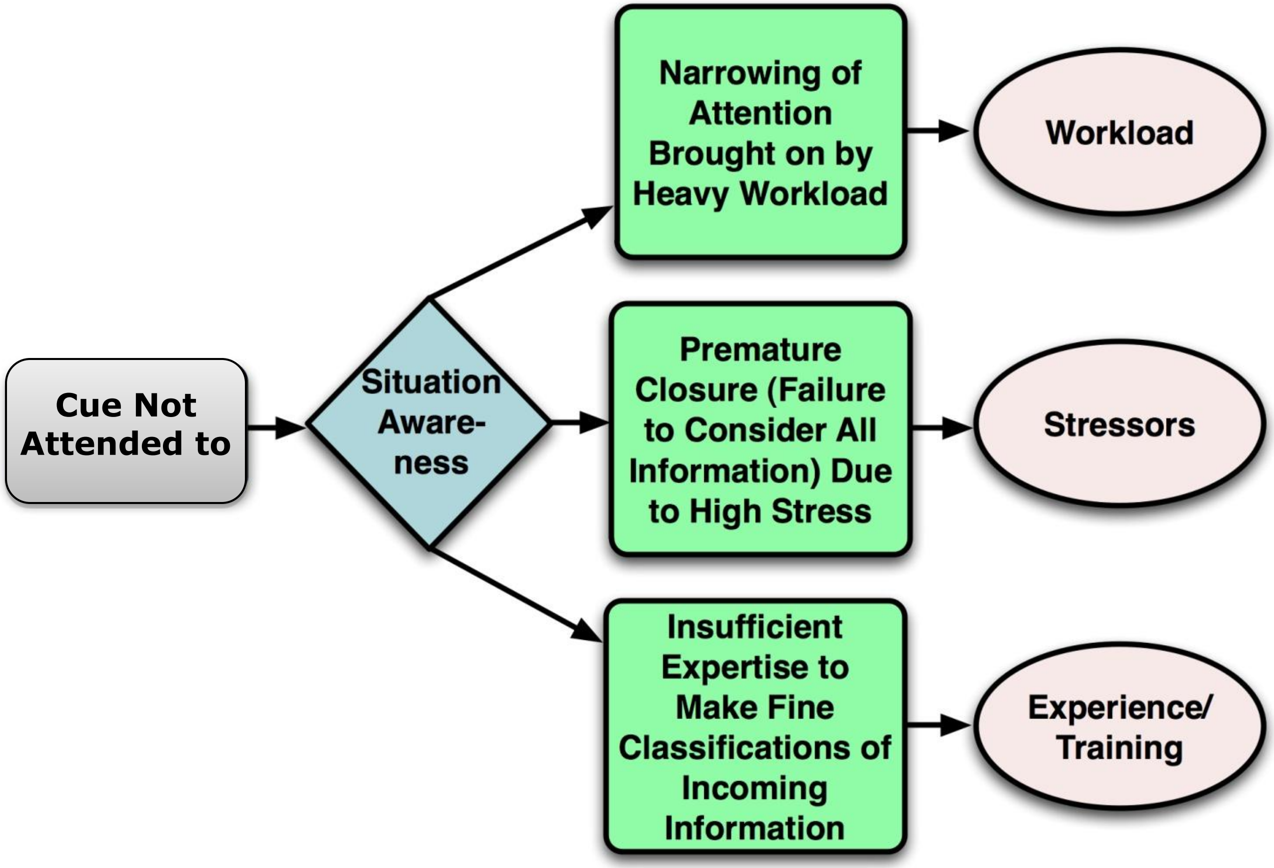
Process Example

Literature Excerpt:	Incorrect information sampling (e.g., inadequate sampling strategy or internal model for directing sampling, misperception of the statistical properties of elements in the environment, forgetting what has already been sampled). [Situation awareness; Endsley, 1995.[4] Error identified by Endsley, page 41, 55.]
Question:	What is the end result of this?
Answer:	• Cues not perceived • Cues not attended to • Cues misread → Proximate Cause

More Context-Specific Instances of Generic Proximate Causes (Examples)

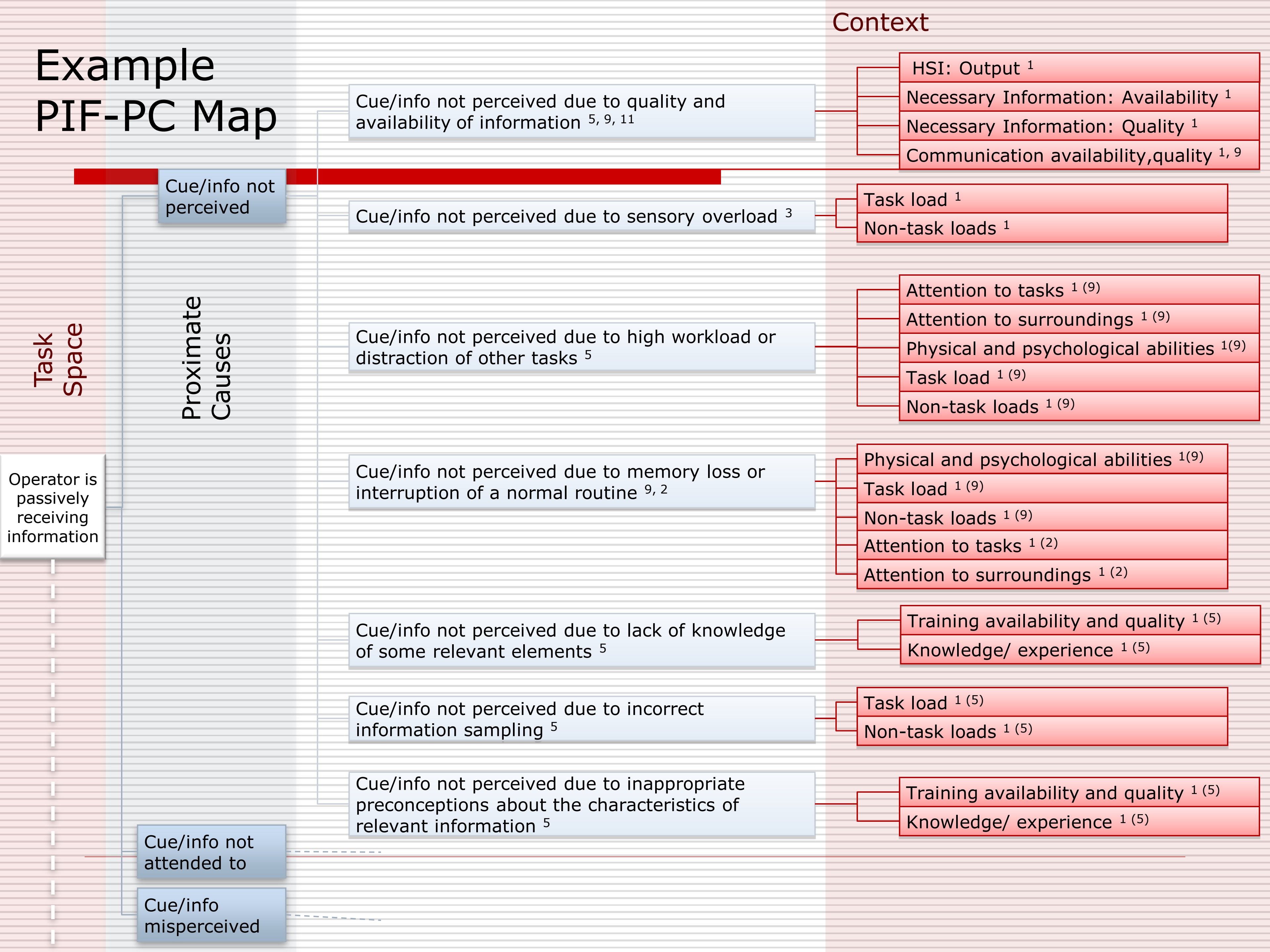
Generic Proximate Cause	Notes
Incomplete/inappropriate list of alternatives (information/explanations/courses of action)	This refers to an incorrect evaluation of alternatives, be it information, diagnoses, or courses of actions. This may include: •Failure to consider all relevant information/explanations/actions •Inappropriately including irrelevant alternatives
Decide upon incorrect alternative (information/explanation/course of action)	This refers to making a choice, often among alternatives. Specific examples that fall into this category of proximate cause include: •Incorrect strategy selected for achieving goal(s) •Decision to skip procedure step(s) •Decision to deviate from procedure •Decision to wait for more information •Decision to take alternate action(s) •Decision to delay action/step(s) •Decision to not collect information

Assignment of PIF to PCs through Psychological Mechanisms



A single proximate cause is explained by different facets of a cognitive model with corresponding different PSFs

Example PIF-PC Map
Task Space
Proximate Causes
Context
Operator is passively receiving information
Cue/info not perceived
Cue/info not attended to
Cue/info misperceived
Cue/info not perceived due to quality and availability of information [5, 9, 11]
Cue/info not perceived due to sensory overload [3]
Cue/info not perceived due to high workload or distraction of other tasks [5]
Cue/info not perceived due to memory loss or interruption of a normal routine [9, 2]
Cue/info not perceived due to lack of knowledge of some relevant elements [5]
Cue/info not perceived due to incorrect information sampling [5]
Cue/info not perceived due to inappropriate preconceptions about the characteristics of relevant information [5]
Cue/info not attended to due to insufficient expertise of knowledge of critical cues in the environment [5]
Cue/info not attended to due to inappropriate preconceptions about the characteristics of relevant information [5]
Cue/info not attended to due to premature closure [5]
Cue/info not attended to due to improper focus of attention, attentional narrowing, or limits to attention sharing ability [5, 6]
Cue/info not attended to due to incorrect information sampling[5]
Cue/info not attended to due to confirmation bias, i.e. evidence does not match the mental model [4]
Cue/info misperceived due to matching to an incorrect template, frame, or mental model [19, 12]
Cue/info misperceived due to improper assemblage of a whole from parts of information [13]
Cue/info misperceived due to improper focus of attention, attentional narrowing, or limits to attention sharing ability[5, 6, 27]
Cue/info misperceived due to inappropriate preconceptions about the characteristics of relevant information [5]
HSI: Output [1]
Necessary Information: Availability [1]
Necessary Information: Quality [1]
Communication availability, quality [1, 9]
Task load [1]
Non-task loads [1]
Physical and psychological abilities [1(9)]
Task load [1 (9)]
Non-task loads [1 (9)]
Attention to tasks [1 (2)]
Attention to surroundings [1 (2)]
Attention to tasks [1 (9)]
Attention to surroundings [1 (9)]
Physical and psychological abilities [1(9)]
Task load [1 (9)]
Non-task loads [1 (9)]
Training availability and quality [1 (5)]
Knowledge/ experience [1 (5)]
Task load [1 (5)]
Non-task loads [1 (5)]
Training availability and quality [1 (5)]
Knowledge/ experience [1 (5)]
Training availability and quality [5]
Knowledge/ experience [5]
Training availability and quality [1 (5)]
Knowledge/ experience [1 (5)]
Task load [1 (5)]
Non-task loads [1 (5)]
Stress [5]
Attention to tasks [1 (5)]
Attention to surroundings [1 (5)]
Physical and psychological abilities [1(5)]
Task load [5]
Non-task loads [5]
Stress [5]
Task load [1 (5)]
Non-task loads [1 (5)]
Training availability and quality [1 (4)]
Knowledge/ experience [1 (4)]
Training availability and quality [1 (4)]
Knowledge/ experience [1 (4)]
HSI: Output [1]
Necessary Information: Quality [1]
Training availability and quality [1]
Knowledge/ experience [1]
Attention to tasks [1 (5)]
Attention to surroundings [1 (5)]
Physical and psychological abilities [1(5)]
Task load [5]
Non-task loads [5]
Stress [5]
Training availability and quality [1 (5)]
Knowledge/ experience [1 (5)]


Context
Example PIF-PC Map
Task Space
Proximate Causes
Operator is passively receiving information
Cue/info not perceived
Cue/info not attended to
Cue/info misperceived
Cue/info not perceived due to quality and availability of information [5, 9, 11]
HSI: Output [1]
Necessary Information: Availability [1]
Necessary Information: Quality [1]
Communication availability,quality [1, 9]
Cue/info not perceived due to sensory overload [3]
Task load [1]
Non-task loads [1]
Cue/info not perceived due to high workload or distraction of other tasks [5]
Attention to tasks [1 (9)]
Attention to surroundings [1 (9)]
Physical and psychological abilities [1(9)]
Task load [1 (9)]
Non-task loads [1 (9)]
Cue/info not perceived due to memory loss or interruption of a normal routine [9, 2]
Physical and psychological abilities [1(9)]
Task load [1 (9)]
Non-task loads [1 (9)]
Attention to tasks [1 (2)]
Attention to surroundings [1 (2)]
Cue/info not perceived due to lack of knowledge of some relevant elements [5]
Training availability and quality [1 (5)]
Knowledge/ experience [1 (5)]
Cue/info not perceived due to incorrect information sampling [5]
Task load [1 (5)]
Non-task loads [1 (5)]
Cue/info not perceived due to inappropriate preconceptions about the characteristics of relevant information [5]
Training availability and quality [1 (5)]
Knowledge/ experience [1 (5)]

Psychological Literature Referenced in Example PC-PIF Assignment

1. Inference
2. Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review, 94*, 115-147.
3. Broadbent, D. (1958). *Perception and Communication*. London: Pergamon Press.
4. Einhorn, H. J. & Hogarth, R. M. (1978). Behavioral decision theory: Processes of judgment and choice. *Annual Review of Psychology, 32*, 53-88.
5. Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37* (1), 32-64.
6. Erickson, C.W., & St. James, J.D. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception and Psychophysics, 40,* 225-240.
7. Henningsen, D. D., Henningsen, M. L. M., Eden, J., & Cruz, M. G. (2006). Examining the symptoms of groupthink and retrospective sensemaking, *Small Group Research*; *37*; (1), 36-64. DOI: 10.1177/1046496405281772.
8. Janis, I.L. (1972). Victims of groupthink: A psychological study of foreign-policy decisions and fiascos. Oxford, England: Houghton Mifflin.
9. Jones, D. G., & Endsley, M. R. (1996). Sources of situation awareness errors in aviation. *Aviation, Space, and Environmnetal Medicine, 67* (6), 507-512.
10. Klein, G. A. (1993). A recognition-primed decision (RPD) model of rapid decision making. In G. A. Klein, J. Orasanu, R. Calderwood, & C. E. Zsambok (Eds.), *Decision Making In Action: Models And Methods*, 138-147. Norwood, NJ: Ablex.
11. Klein, G. A. (1998). *Sources of power: How people make decisions*. Cambridge, MA: MIT Press.
12. Klein, G., & Moon, B. (2006). Making sense of sensemaking 2: A macrocognitive model. *IEEE Intelligent Systems, 21* (5), 88-91.
13. Köhler, W. (1947). *Gestalt Psychology: An Introduction to New Concepts in Modern Psychology*. New York: Liveright.
14. Letsky, M. P., Warner, N. W., Fiore, S. M., & Smith, C. A. P. (2008). *Macrocognition in Teams: Theories and Methodologies*. Hampshire, England: Ashgate Publishing Limited.
15. Lipshitz, R. (1993). Converging themes in the study of decision making in realistic settings. In G. A. Klein, J. Orasanu, R. Calderwood, & C. E. Zsambok (Eds*.), Decision making in action: Models and methods*, 103-137. Norwood, NJ: Ablex.
16. Massaro, D.W., & Cohen, M.M. (1991). Integration versus interactive activation: The joint influence of stimulus and context in perception. *Cognitive Psychology, 23*, 558-614.
17. Neisser, U. (1967). *Cognitive Psychology*. New York: Appleton.
18. Orasanu, J., & Martin, L. (1998). Errors in aviation decision making: A factor in accidents and incidents. In *Proceedings of the Workshop on Human Error, Safety, and Systems Development,* Seattle, WA.
19. Phillips, W.A. (1974). On the distinction between sensory storage and short term visual memory. *Perception and Psychophysics, 16*, 283-290.
20. Proctor, R. W. & Van Zandt, T. (2008). *Human Factors in Simple and Complex Systems, 2nd Ed*. Taylor & Francis.
21. Reason, J. & Hobbs, A. (2003). *Managing Maintenance Error: A Practical Guide*. Ashgate Publishing Limited.
22. Reason, J. (2008). The Human Contribution: Unsafe Acts, Accidents, and Heroic Recoveries. Ashgate Publishing Limited.
23. Reicher,G. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology, 81*, 275-280.
24. Roth, E.M. (1997). Analysis of decision making in nuclear power plant emergencies: An investigation of aided decision making. In C.E. Zsambok & G. Klein (Eds.), *Naturalistic Decision Making*, 175-182. Mahwah, NJ: Lawrence Erlbaum.
25. Siddle, B. K. (2008). “The Stress Paradox” (p.28-32), in the War(x) on Trauma: Lesson Learned from a Decade of Conflict, Elsevier. Posted as online article at http://www.jems.com/Images/nar_supp32_new_tcm16-171092.pdf.
26. Sternberg, R. J., & Davidson, J. E. (Eds.) (1984). Conceptions of giftedness. New York: Cambridge University Press.
27. Treisman, A. M. (1991). Search, similarity, and integration of features between and within dimensions. Journal of Experimental Psychology: Human Perception and Performance, 17, 652-676.
28. Tversky, A. & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1124-1131.
29. Tversky, A. & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211*, 1124-1131.
30. Warner, N. W., & Letsky, M. P. (2008). Empirical Model of Team Collaboration Focus on Macrocognition. In M. P., Letsky, N. W. Warner, S. M. Fiore, & C. A. P. Smith (Eds.), *Macrocognition in Teams: Theories and Methodologies*, 15-33. Hampshire, England: Ashgate Publishing Limited.
31. Warren, R.M. (1970). Perceptual restorations of missing speech sounds. *Science, 167*, 392-393.
32. Wickens, C. D., Lee, J. D., Liu, Y. & Becker, S. E. G. (2004). *An Introduction to Human Factors Engineering (2nd Ed)*. Pearson Prentice Hall

Status

- Psychological literature nearly completed
- Taxonomy of PIFs completed
- Mapping of generic PCs to PIFs nearly completed
- Identification of specific instances of generic PCs (more relevant to control room operations) nearly completed
- Establishing guidance for use of the model for Qualitative and Quantitative Analysis
    - Parts already in place

Status

- We are developing a logic structure to complete the causal model and support quantitative approach
  - Relation between PCs and HFE (in part to treat “local recovery” of PCs)
  - Treating interdependencies of PCs due to common PIFs
  - Treating PIF interdependencies
  - Potentially a subject for next ACRS Subcommittee meeting
- Coordination with other projects to examine absorbability of PCs in (a) operating experience in coordination with HRA data base development activities, and (b) simulator data collection and analysis

References

1. Smidts, C., Shen, S.-H., & Mosleh, A. (1997). The IDA cognitive model for the analysis of nuclear power plant operator response under accident condition. Part I: problem solving and decision making model. *Reliability Engineering & System Safety, 55,* 51-71.
2. Baddeley, A. (2007). *Working Memory, Thought, and Action.* Oxford, UK: Oxford University Press.
3. Chang, Y. H. J., & Mosleh, IDAC Methdology (5 Papers Published in RESS)
4. Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37* (1), 32-64.
5. Broadbent, D. (1958). *Perception and Communication*. London: Pergamon Press.
6. Neisser, U. (1967). *Cognitive Psychology*. New York: Appleton.
7. Jones, D. G., & Endsley, M. R. (1996). Sources of situation awareness errors in aviation. *Aviation, Space, and Environmental Medicine, 67* (6), 507-512.
8. Roth, E.M. (1997). Analysis of decision making in nuclear power plant emergencies: An investigation of aided decision making. In C.E. Zsambok & G. Klein (Eds.), *Naturalistic Decision Making*, (pp. 175-182). Mahwah, NJ: Lawrence Erlbaum.
9. Klein, G., & Moon, B. (2006). Making sense of sensemaking 2: A macrocognitive model. *IEEE Intelligent Systems, 21* (5), 88-91.
10. Klein, G. A. (1993). A recognition-primed decision (RPD) model of rapid decision making. In G. A. Klein, J. Orasanu, R. Calderwood, & C. E. Zsambok (Eds.), *Decision Making In Action: Models And Methods* (pp. 138-147). Norwood, NJ: Ablex.
11. Lipshitz, R. (1993). Converging themes in the study of decision making in realistic settings. In G. A. Klein, J. Orasanu, R. Calderwood, & C. E. Zsambok (Eds.), *Decision Making In Action: Models And Methods* (pp. 103-137). Norwood, NJ: Ablex.
12. Reason, J. (2008). *The Human Contribution: Unsafe Acts, Accidents, and Heroic Recoveries*. Ashgate Publishing Limited.
13. Reason, J. & Hobbs, A. (2003). *Managing Maintenance Error: A Practical Guide*. Ashgate Publishing Limited.
14. Groth, K., & Mosleh, A. (2009). *A Data-informed Model Of Performance Shaping Factors For Use In Human Reliability Analysis*. Doctoral Dissertation, University of Maryland, Center for Risk and Reliability, College Park, MD.

Back-Up Slides

Main Cognitive Framework

Information Processing Framework

- Adopted as the overarching framework
- Represents generalization of multiple cognitive models
- Model is used in existing HRA methods
- 3 stage human information processing model
  - Includes all relevant aspects of human performance
  - Maps to more complex models
  - Offers More distinct lines based on cognitive demands
- Not intended to be used by practitioners

I-D-A Stages

- "I" includes (a) sensing and (b) perceiving information.
- Information perception in this stage includes:
  - assigning meaning in a generic sense (e.g., observed color is red),
  - generic, not context-specific "light chunking" of the information (e.g., flashing light plus siren=alarm)
  - initial classification and grouping of incoming information.
- It stops short of further inference and conclusions based on processed information as a whole.

I-D-A Stages

- "D" covers the operator response phases of "situation assessment / diagnosis", and "response planning." It contains problem solving through goal setting, selection of and approach to achieve the goal(s), and corresponding decisions*

- "A" refers to the execution of the actions decided upon. The actions are skill-based, and of relatively low cognitive complexity

* The complexity and exact nature of the activities within "D" depends on the levels of task decomposition

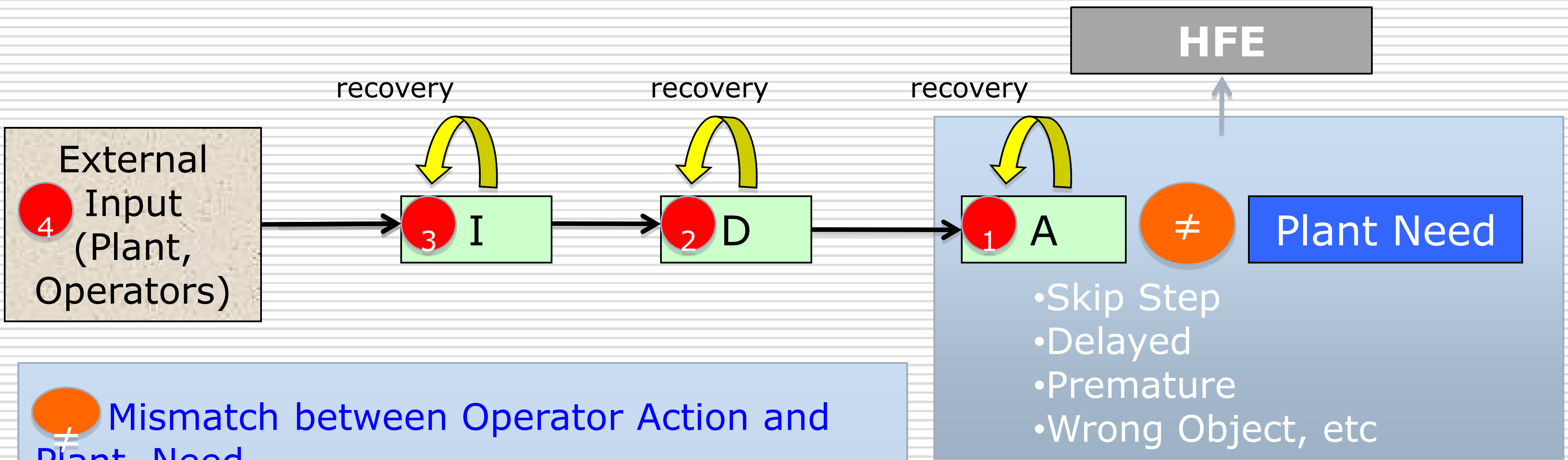
Information Processing Framework

- Why is using an information processing framework (i.e., I-D-A[1]) appropriate?
  - Simple model that covers all aspects of human performance of interest for HRA
  - Aids in model development by guiding literature review to identify relevant areas of psychology
- IDA is useful both as way of "tasks decomposition" as well as an overarching simple model of cognitive response
- With some modification the model is applicable irrespective of whether "unit of analysis" is *crew* or an *individual operator*

Causal Model of Error

Human Error in IDA Framework

- Error defined in terms of the operator failing to meet a plant need. Cause of errors is traced through the IDA phases



≠ Mismatch between Operator Action and Plant Need

1 Failure of execution of action given correct decision

2 Failure of decision given correct required information

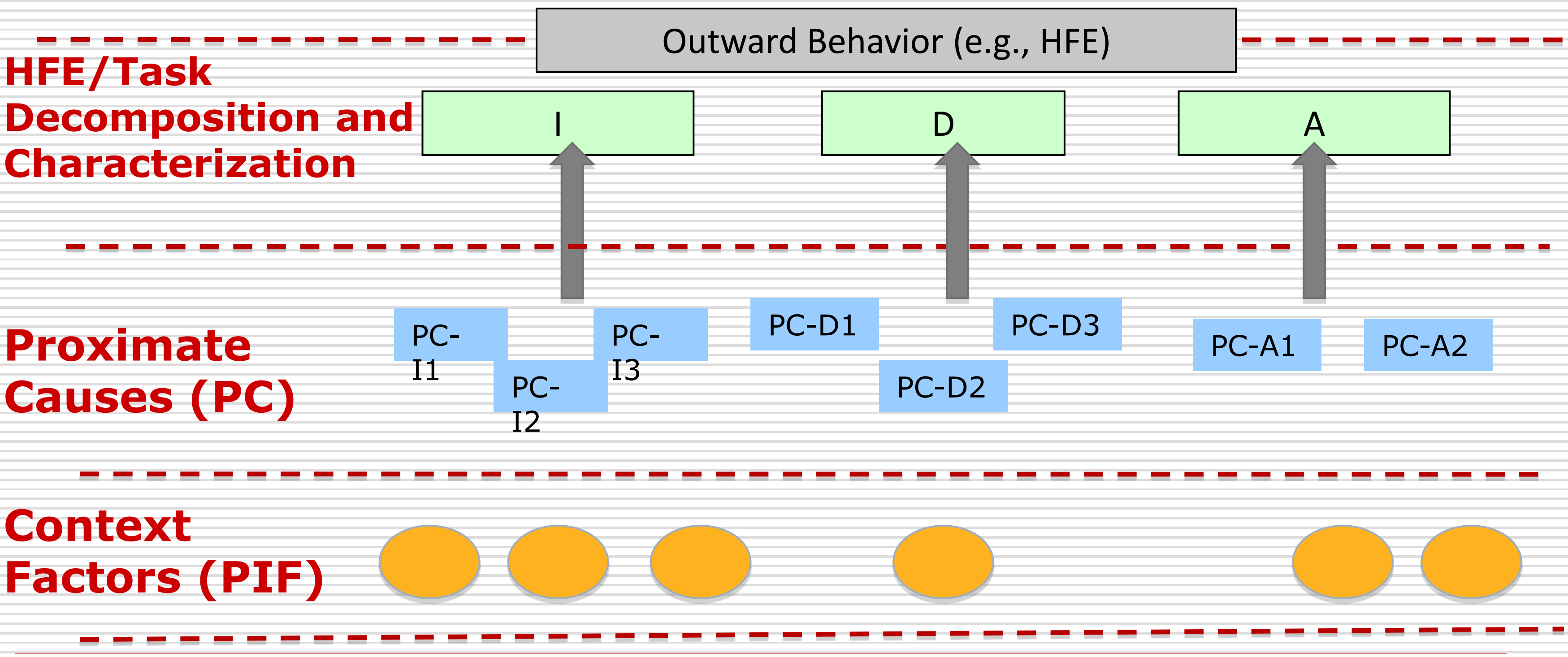
3 Failure of information perception given availability of information

4 Incorrect information from external source

Logic of Failure
HFE = I+D+A
(logical OR)

Simplified Causal Model

- A simple three-tiered causal model is used to represent the path from context to outward behavior (e.g., a HFE as defined in PRA)



Model Structure

- The model distinguishes between:
  - Task Space
    - used as a task decomposition to describe the type of activity in the event response
  - Cognitive Space (Macrocognitive Functions):
    - describe the cognitive activity related to the task
- The macrocognitive function is always assessed in relation to the task
  - Cognition occurs in context of the task
  - Any or all of the macrocognitive functions can occur for each of the task types

Examples Frameworks from Literature

Naturalistic Decision-Making (NDM)

- Attempts to model how people make decisions in familiar, real world contexts
- Decisions characterized by:
  - Uncertain, dynamic environments
  - Ill-defined problem and goals
  - Action feedback (previous decisions affect future options)
  - Time stress
  - High stakes
  - Multiple players

Example Frameworks from Literature

Recognition Primed Decision-Making

- NDM strategy based on decision-making behavior of experienced firefighters
  - Rarely did decision-makers consider even two options concurrently
  - A search for an optimal choice could delay action long enough to result in loss of situational control
- General strategy is to match current situation to a prototypical situation drawn from experience
- Options considered sequentially

Emphasizes the need to develop an accurate situational assessment (information perception is important...)

Example Frameworks from Literature

- "Cognitive Engineering" Perspective [8]
  - Human behavior is goal-driven, actions are selected in support of a goal, …. involves iterative cognitive cycles
  - Human behavior may be better modeled as 'multi-level' nested goal-driven actions.
- IDAC Model [3]
  - Operator "problem solving" behavior is
    - Rational (has an explanation),
    - Goal-Oriented (single/ multiple, parallel/sequential)
    - Use identifiable "problem-solving strategies"
    - Dynamic
  - Operator "decision making" (choice among alternatives) is
    - Rational (has an explanation)
    - Risk-Benefit Based

More on "D" Model

- Possible crew goals are context/scenario-specific

- Most-likely problem solving strategy in Full Power scenarios
  - Knowledge-Supplemented Procedural Response

- All problem solving and decision making strategies identified for NPP crew response can be constructed in terms of Reason's
  - Basic Similarity Matching (BSM)
  - Frequency Gambling (FG)

Back up on Results

"I" Proximate Causes of Failure

Proximate Causes	Passive (Receiving I)	Active (Seeking I)
Cues/Inf Not Sensed	X	_
Cues/Inf Not Perceived	X	_
Cues/Inf Not Attended To	X	_
Cues/Inf Misperceived	X	X
Cues/Inf Discounted/ Dismissed (as irrelevant or unimportant)	X	X
Wrong Cues/Inf Attended To	X	X
Internationally Not Gathered	_	X

"D" Model and Proximate Causes

- "D" Covers Two (Cognitively) Distinct Activities
  - **Problem Solving** (e.g., Diagnosis, Develop Action Plan)
  - **Decision Making** – making choice among alternatives (The problem-solving process involves a series of decisions to be made or solutions to be selected based on available alternatives)
    - Follow EOPs vs. CSF-procedures
    - Restoring AFW flow vs. performing Feed & Bleed in Loss of Heat Sink scenarios
- Possibility subject to different influencing factors

"D" Generic Proximate Causes

Generic Proximate Cause	Notes/Mechanisms
Incorrect understanding/diagnosis of the situation	• Cues/info misunderstood / misinterpreted ~ Single piece of information ~ synthesizing multiple pieces of information • Incorrect/incomplete mental model • Incorrect/incomplete projection of future status/ mental simulation • Incorrect/inappropriate mapping of information/cues to a mental model • Failure to change incorrect mental model • Inappropriate change of mental model • Failure to deviate from or change procedures when the situation warrants it • Inappropriate deviation from or change of procedures (these last two items may fit better elsewhere)
Inability to develop diagnosis from available	•No mental model exists that matches the information, no pattern with which to interpret the data

"D" Generic Proximate Causes

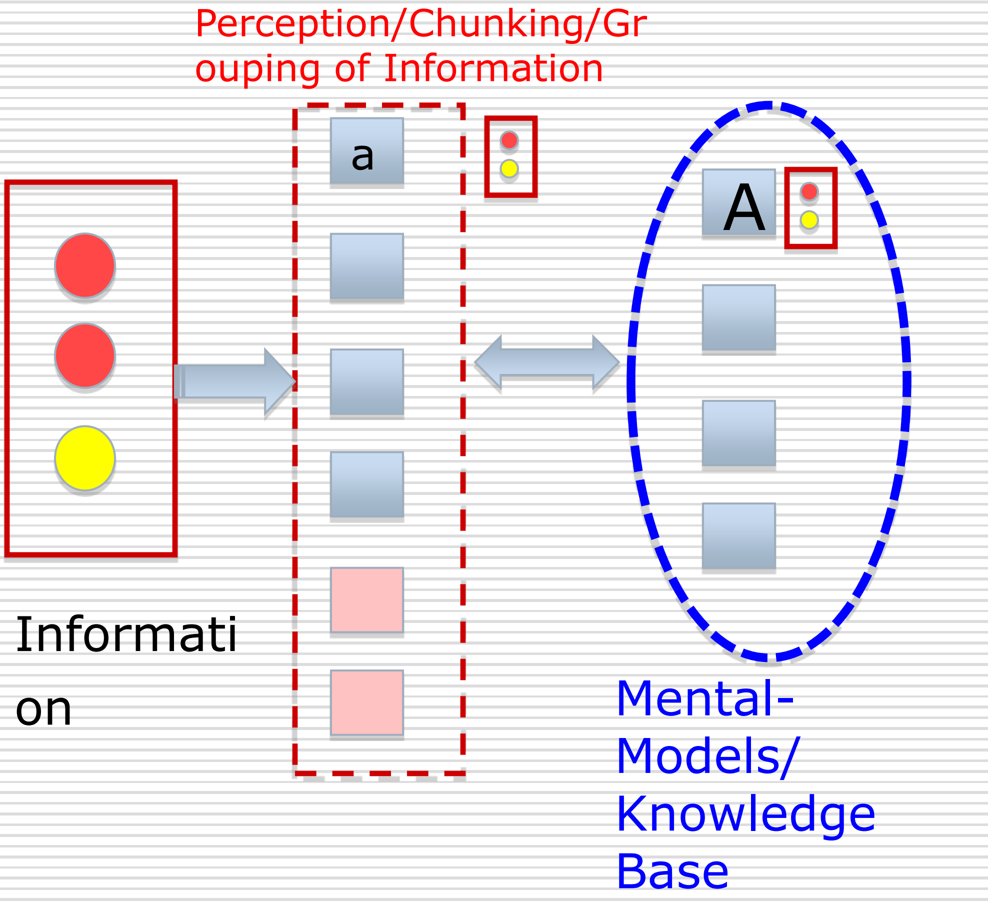
Generic Proximate Cause	Notes/Mechanisms
Inappropriate goal/priority selected	Specifically, in using the term *goal,* we're referring to *functional goals* (i.e., plant functions, such as establishing AFW). Reasons for not selecting the correct functional goal may include: -- Administrative goals/rules/policies -- Cost/benefit/risk/safety prioritization -- Personal goals (e.g., don't be wrong, avoid embarrassment, ego issues, competition, etc.)
Failure to properly maintain/balance multiple goals/priorities	Involves assigning improper priorities to multiple goals or tasks.

“D” Generic Proximate Causes

Generic Proximate Cause	Notes/Mechanisms
Consideration of incomplete/inappropriate list of alternatives (information/explanations/courses of action)	This refers to an incorrect evaluation of alternatives, be it information, diagnoses, or courses of actions. This may include: •Failure to consider all relevant information/explanations/actions •Inappropriately including irrelevant alternatives
Decide upon incorrect alternative (information/explanation/course of action)	This refers to making a choice, often among alternatives. Specific examples that fall into this category of proximate cause include: •Incorrect strategy selected for achieving goal(s) •Decision to skip procedure step(s) •Decision to deviate from procedure •Decision to wait for more information •Decision to take alternate action(s) •Decision to delay action/step(s) •Decision to not collect information

Pattern Matching (Similarity Matching) Cognitive Mechanism

- Pattern Matching is a prominent mode of cognition in much of the literature
  - Leads to operators forming mental models/beliefs
  - Judged to be applicable to "Knowledge-supplemented procedure following" strategy



Use of BSM in identifying failure mechanisms

- This characterization helps to identify some key failure mechanisms.

  - Failure of BSM due to memory recall failure: operator does not recall the memorized match to cue
  - Failure of BSM due to poor training: operator's knowledge base (from training) lacks match to cue
  - Failure of BSM due to regency of a cue-situation match (failure because current situation is different)

- Such rules apply to I, D, and A elements

I-D-A Generic Proximate Causes

I:

- Cues/info not sensed
- Cues/info not perceived
- Cues/info not attended to
- Cues/info misperceived
- Cues/info discounted/dismissed
- Wrong cues/info attended to
- **D:**
- Incorrect understanding/diagnosis of the situation
- Inability to develop diagnosis from available info (no diagnosis, failure to diagnose)
- Inappropriate goal/priority selected
- Failure to properly maintain/balance multiple goals/priorities
- Consideration of incomplete/inappropriate list of alternatives (information/explanations/courses of action)
- Decide upon incorrect alternative (information/explanation/course of action

A:

- Omission (Failure to perform action)
- Incorrect execution (force, direction, degree/distance, object)
- Incorrect timing
- Incorrect order/sequence
- Commit incorrect action (specifically an accidental or unintentional incorrect action, e.g., habit intrusion, interference error, perseverations)
- Failure to communicate information verbally or nonverbally
- Incorrectly communicate information verbally or nonverbally

PSF Taxonomy and Properties

Organizational PSFs

- Training Program
  - Availability
  - Quality
- Corrective Action Program
  - Availability
  - Quality
- Other Programs
  - Availability
  - Quality
- Safety Culture
- Management Activities
  - Staffing
    - *Number*
    - *Qualifications*
    - *Team composition*
  - Scheduling
    - *Prioritization*
    - *Frequency*
- Workplace adequacy
- Resources
  - Procedures
    - Availability
    - Quality
  - Tools
    - Availability
    - Quality
  - Necessary Information
    - Availability
    - Quality

Team PSFs

- Communication
  - Availability
  - Quality
- Direct Supervision
  - Leadership
  - Team member
- Team Coordination
- Team Cohesion
- Role Awareness

Personal PSFs

- Attention
  - To Task
  - To Surroundings
- Physical & Psychological Abilities
  - Alertness
  - Fatigue
  - Impairment
  - Sensory Limits
  - Physical attributes
  - Other
- Morale/Motivation/Attitude (MMA)
  - *Problem Solving Style*
  - *Information Use*
  - *Prioritization*
    - *Conflicting Goals*
    - *Task Order*
  - *Compliance*
- Knowledge/Experience
- Skills
- Familiarity with Situation
- Bias

Situation PSFs

- External Environment
- Hardware & Software Conditioning Events
- Task Load
- Time Load
- Other Loads
  - Non-task
  - Passive Information
- Task Complexity
  - Cognitive
  - Task Execution
- Perceived Situation:
  - Severity
  - Urgency
- Perceived Decision:
  - Responsibility
  - Impact
    - Personal
    - Plant
    - Society

Machine Design PSFs

- HSI
  - Input
  - Output
- System Responses
  - *Ambiguity*

Italicized elements are behaviors or metrics associated with the parent PIF.

Generic Proximate Causes Identified

- Cues/info not sensed
- Cues/info not perceived
- Cues/info not attended to
- Cues/info misperceived
- Cues/info discounted/dismissed
- Wrong cues/info attended to

- Incorrect understanding/diagnosis of the situation
- Inability to develop diagnosis from available info (no diagnosis, failure to diagnose)
- Inappropriate goal/priority selected
- Failure to properly maintain/balance multiple goals/priorities
- Consideration of incomplete/inappropriate list of alternatives (information/explanations/courses of action)
- Decide upon incorrect alternative (information/explanation/course of action

- Omission (Failure to perform action)
- Incorrect execution (force, direction, degree/distance, object)
- Incorrect timing
- Incorrect order/sequence
- Commit incorrect action (specifically an accidental or unintentional incorrect action, e.g., habit intrusion, interference error, perseverations)
- Failure to communicate information verbally or nonverbally
- Incorrectly communicate information verbally or nonverbally

Developing an HRA Quantification Model

Gareth W Parry
ERIN Engineering and Research Inc.
gwparry47@erineng.com

ACRS PRA Subcommittee
Workshop

October 18, 2010

What is the HRA Quantification Model?

- For the purposes of this presentation, the HRA quantification model is the tool that generates the human error probability (HEP) associated with a human failure event (HFE)
- The model will use as input the performance influencing factors(PIFs) that define the overall context for the HFE
  - Plant conditions
  - PSFs

Groundrules for Construction of HRA Model

- The model should:
  - 1. have a sound basis consistent with the cognitive psychology and behavioral science disciplines
  - 2. be practical and applicable directly to the HFEs defined in the PRA.
  - 3. provide the user with the tools to identify the critical elements of context (PIFs) used for the quantification of HEPs
  - 4. facilitate intra and inter analyst consistency and reliability

Theoretical Basis for the Quantifiaction Model

- Nested Information-Decision-Action (IDA) framework
- Set of proximate causes for human failure in the NPP environment
- Cognitive mechanisms linking PIFs to proximate causes

Approach

- For each proximate cause, construct a decision tree:
  - The headers on the tree represent critical PIFs (groundrule 1)
  - For each branch, a set of questions is provided to determine as objectively as possible the existence or not of the PIF (groundrule 3)
  - Each path through the tree represents an explanation of the human failure in terms of a cognitive mechanism and the PIFs (groundrule 1)

Approach (Cont’d)

- Use expert judgment (SRM team)to associate an HEP with each path through the tree
- The PIFs are not independent - they may reinforce one another or counteract one another. There is no empirical basis to guide the experts other than their general understanding
- Proposed approach. For each decision tree:
  - rank the paths by HEP
  - reach consensus on two bounding case HEPs
  - Interpolate

Quantification Model

- The intent is that the model will be a consensus model, thus satisfying groundrules 2 and 4
- Use:
  - defines the required qualitative analysis to complement that provided in the definition of the HFEs
  - can be used for existing PRAs or to support an ATHEANA type analysis of deviation scenarios
  - can provide guidance on choice of subcontexts to explore deviation scenarios
  - could be used iteratively with a CRT or ATHEANA approach for developing a more detailed PRA/HRA

Quantification

- For a scenario S with an associated context

$$HEP(HFE|S) = \sum_i Prob\ (\mathrm{PC_i}|S)$$

- For an HFE with the context broken into subcontexts

$$HEP(HFE|S) = \sum_j \sum_i Prob(\mathrm{PC_i}|\mathrm{S_j}).Prob(\mathrm{Sj}|S)$$

Dependency

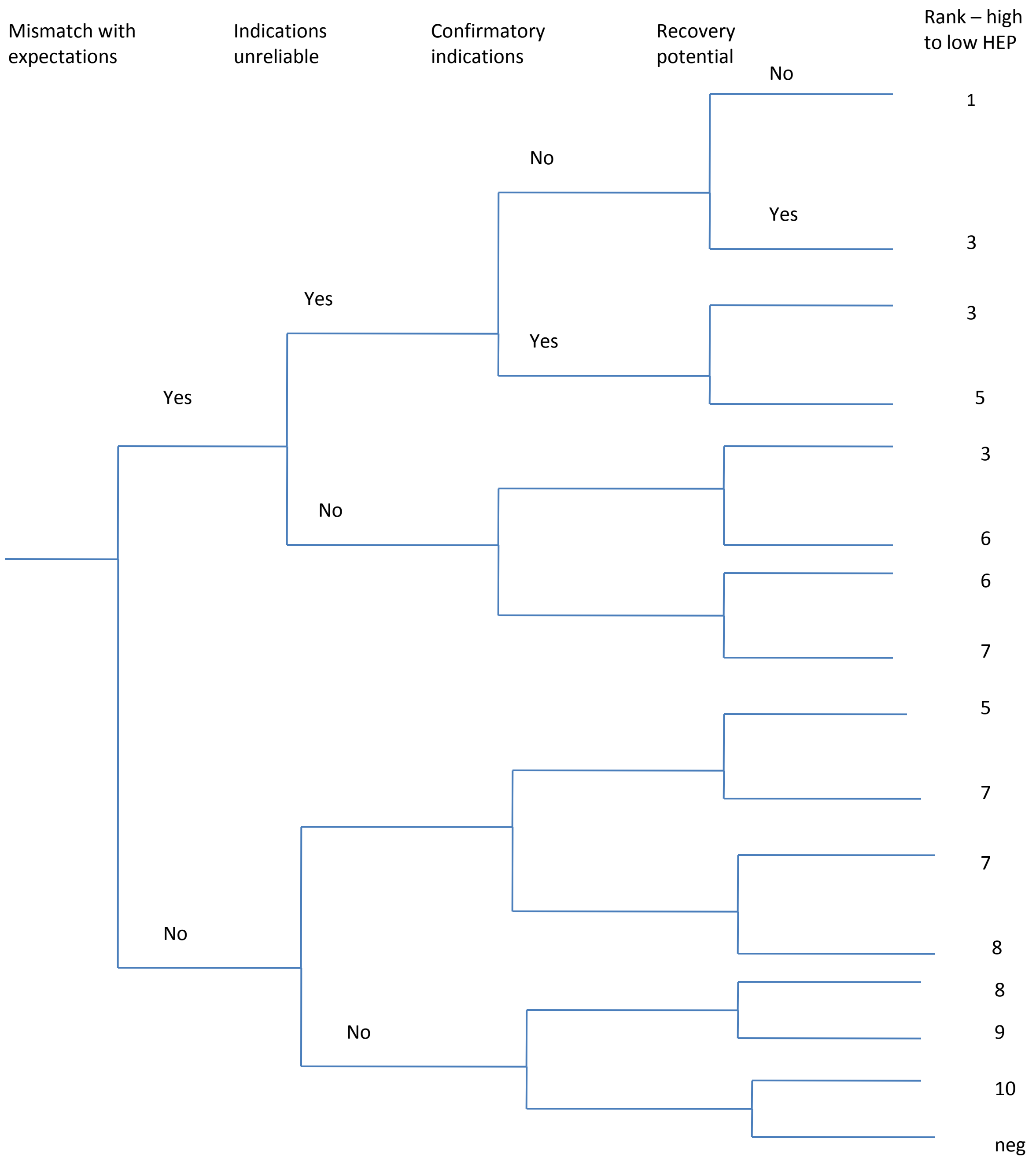
- The focus on causes provides a more rational basis for addressing dependency between HFEs occurring on the same accident sequence cut set.

Back-up Slides

Example of a decision tree for the proximate cause: Intentionally dismiss information

PIFs

- Bias (incl. from training)
  - Confirmation of theory/preconceived notion
  - Against a specific information source
    - (e.g., they think the indicator is wrong because of previous experience, or because maintenance has repeatedly tried to fix the indicator, etc)
- Conditioning Events
  - Past experience with specific indicator
- Ambiguity of system responses (this is one of main driver for this proximate cause)
  - Must be an ambiguous situation or they don’t have the option to discount data
- Training,
- Knowledge & Experience
  - Failure to recognize the relevance of the additional data
  - Misperceive the utility of the information
  - Misranking of solution alternatives


Mismatch with expectations
Indications unreliable
Confirmatory indications
Recovery potential
Rank – high to low HEP
No
1
No
Yes
3
Yes
3
Yes
Yes
5
3
No
6
6
7
5
7
7
No
8
8
9
No
10
neg

Example questions

- BP3: Confirmatory information – can the significance of the missing information be validated by other information?
  - If this is possible, this provides a defense against dismissing the information. Note that this is a procedure/knowledge driven PIF.
  - Are there additional indications that would typically be used to confirm the plant status indicated by the information (e.g., pump amps to confirm a pump unavailable)?
  - Is checking these additional sources emphasized in training?
- If the answer is YES to both these questions, there should be a lower likelihood of dismissing the information.